M.M. POSTNIKOV

LEÇONS DE GÉOMÉTRIE

5e semestre

GROUPES ET ALGÈBRES DE LIE

M. POSTNIKOV

LEÇONS DE GÉOMÉTRIE

GROUPES ET ALGÈBRES DE LIE

ÉDITIONS MIR · MOSCOU

TABLE DES MATIÈRES

AVANT-PROPOS

La théorie des groupes de Lie est basée sur le théorème de Cartan d'équivalence des catégories de groupes de Lie simplement connexes et des catégories d'algèbres de Lie. Cet ouvrage vise à la démonstration de ce théorème et des principaux faits qui y sont rattachés. Nous glisserons donc sur des résultats plus profonds découlant de ce théorème. De même la théorie des algèbres de Lie n'est exposée que dans la mesure nécessaire à la démonstration du théorème de Cartan.

Ce cours spécial semestriel de 21 leçons prolonge une série *) qui est la transcription presque intégrale des leçons faites par l'auteur aux élèves du second (et du troisième) cycle de la faculté de mécanique et de mathématiques de l'Université de Moscou. Mais il diffère des livres I et II qui, eux, reproduisaient un cours obligatoire.

La destination de ce cours aux élèves de quatrième et de cinquième année et aux boursiers de thèse a permis en deux heures académiques (90 min) d'aborder plus de sujets que dans les leçons qui visaient les élèves de première année. En supprimant la récréation et en « mordant » sur les horaires du cours suivant, l'auteur a réussi à faire de deux heures académiques deux heures « astronomiques » (120 min) et ainsi à doubler pratiquement le volume de chaque leçon. Certes, avec un programme moins chargé — disons un cours étalé sur une année et non sur un semestre — chacune de nos leçons aurait recouvert pratiquement une leçon et demie ou deux ordinaires. Aussi serait-il plus logique de voir ici un cours spécial d'une année (mais qui a tout de même été fait en un semestre dans les conditions signalées ci-dessus).

La pénurie de temps nous a contraint à nous limiter souvent aux seules idées des démonstrations en abandonnant les détails au lecteur. Les assertions auxiliaires des autres disciplines mathématiques ont été simplement formulées avec les références nécessaires, et les exemples illustrant la théorie générale, simplement décrits, leur étude détaillée ayant été laissée au soin du lecteur.

---

*) Voir M. Postnikov, *Leçons de géométrie*. Ier semestre. Géométrie analytique. Traduction française Editions Mir, 1981, et *Leçons de géométrie*. IIe semestre. Algèbre linéaire et géométrie différentielle. Traduction française Editions Mir, 1981 (sont désignés dans les références par I et II suivis du numéro de la leçon). Les semestres III et IV sont *in statu nascendi*.

En couchant ces leçons par écrit point n'est besoin de respecter ces particularités, bien plus, toutes les démonstrations doivent être conduites soigneusement, les exemples étudiés complètement et les lemmes « auxiliaires » prouvés. Ceci aura pour effet de doubler ou de tripler le volume des leçons.

Tout professeur, même s'il suppose à ses élèves un certain niveau de connaissance, est contraint de rappeler, ne fût-ce que brièvement, les notions préliminaires fondamentales. Par écrit, ces rappels occupent une place assez importante. Ceci explique l'ampleur de certaines leçons. En fait, chaque leçon de cet ouvrage est la transcription d'une leçon orale (aux artifices près de l'auteur!).

Les leçons ont été partagées en cinq cycles. Dans le premier (leçons 1, 2 et 3) sont introduites et illustrées par des exemples les notions fondamentales de groupe de Lie, d'algèbre de Lie d'un groupe de Lie.

Le deuxième cycle (leçons 4, 5, 6 et 7) est consacré à la « théorie locale » des groupes de Lie. Les leçons 4 et 6 établissent l'équivalence des catégories d'algèbres de Lie et de groupuscules analytiques (= groupes locaux) de Lie. L'outil mathématique nécessaire est développé à la leçon 5. Dans la leçon 7 on montre que la condition d'analyticité ne restreint pas la généralité. On y étudie aussi les sous-groupuscules et les groupuscules quotients.

Les leçons 8, 9 et 10 sont consacrées à la globalisation de la théorie. La leçon 8 traite de la théorie des revêtements (au sens de Chevalley, c'est-à-dire « sans chemins »). Dans la leçon 9, on construit un groupe de revêtement universel; dans la leçon 10, on formule et on discute le théorème de Cartan. Ce théorème n'est pas prouvé mais ramené simplement au théorème d'Ado d'existence de la représentation linéaire exacte de toute algèbre de Lie.

Ces trois cycles peuvent constituer la matière d'un petit cours d'initiation à la théorie des groupes de Lie.

Dans les leçons 11 et 12 sont étudiés les sous-groupes et les groupes quotients des groupes de Lie. La leçon 13 est consacrée aux algèbres de Clifford et aux groupes de spineurs. Dans les leçons 14, 15 et 16 on se penche en détail sur les groupes de Lie G_2 et F_4 et sur l'appareil mathématique *ad hoc*.

Les dernières leçons 17 à 21 sont purement algébriques et sont pratiquement indépendantes des autres (hormis la leçon 20 qui « fait bande à part»). Formellement elles sont consacrées à la démonstration du théorème d'Ado, mais en réalité elles contiennent une partie assez importante de la théorie des algèbres de Lie (critère de Cartan de résolubilité et de semi-simplicité, lemmes de Whitehead, théorèmes de Weyl et de Lévi) qui présente un intérêt en soi.

En conclusion je voudrais exprimer ma gratitude à V. Popov dont la contribution au perfectionnement du manuscrit dépasse de loin les obligations habituelles du rédacteur.

M. Postnikov

LEÇON 1

Groupes topologiques et différentiables.— Affaiblissement des conditions définissant les groupes de Lie.— Exemples de groupes de Lie.— Transformation de Cayley.— Autres exemples de groupes de Lie.— Espaces et groupes connexes et connexes par arcs.— Réduction de groupes différentiables à des groupes connexes.— Exemples de groupes de Lie connexes.

Soit G un groupe muni d'une structure de variété différentiable *).

Définition 1. Un groupe G est un *groupe de Lie* (ou un *groupe différentiable*) si les applications

(1) $$G \times G \to G, \quad (a, b) \mapsto ab,$$

(2) $$G \to G, \quad a \mapsto a^{-1},$$

sont différentiables.

Soient G et H des groupes de Lie. Une application $G \to H$ est par définition un *morphisme* de groupes de Lie (ou un *homomorphisme différentiable* de ces groupes) si elle est un homomorphisme pour la structure abstraite de ces groupes et une application différentiable pour la structure de variété de ces groupes.

Il est clair que tous les groupes de Lie et leurs homomorphismes forment une catégorie que nous désignerons par GR-DIFF.

Remarque 1. Il existe à proprement parler une famille dénombrable de catégories GR-DIFF selon la classe C^r (où soit $r \in [2, \infty]$, soit $r = \omega$) de différentiabilité exigée des variétés envisagées. Mais, en réalité, rien ne dépend de r, puisque, comme nous le verrons dans

---

*) Nous supposons au lecteur les notions fondamentales de la théorie des variétés différentiables.

Les notions fondamentales de la théorie des groupes sont accessibles par exemple dans l'ouvrage de A. Kostrikin *Introduction à l'algèbre*. Editions Mir, 1982.

la leçon 7, tout groupe C^r-différentiable est C^r-isomorphe à un groupe analytique (de classe C^ω).

Remarque 2. Certains auteurs font une subtile distinction entre les groupes de Lie et les groupes différentiables (notamment analytiques). Pour notre part nous considérerons que ces deux termes sont synonymes en accordant toutefois la préférence au premier, car plus courant et plus traditionnel.

De façon analogue, un groupe G muni d'une structure d'espace topologique est un *groupe topologique* si les applications (1) et (2) sont continues. On dit qu'un homomorphisme $G \to H$ de groupes topologiques est *continu* s'il est une application continue. Les groupes topologiques et leurs homomorphismes continus forment la catégorie GR-TOP.

On rappelle qu'un espace topologique M est *séparé* si deux quelconques de ses points possèdent des voisinages disjoints, c'est-à-dire si la diagonale Δ (sous-ensemble du produit $M \times M$ composé des points de la forme (x, x), $x \in M$) est fermée dans M. En général, nous n'exigerons pas que les groupes topologiques (de même que les variétés différentiables) soient séparés.

Lemme 1. *Une condition nécessaire et suffisante pour qu'un groupe G soit séparé est que son unité soit fermée.*

D é m o n s t r a t i o n. Dans tout espace séparé, l'unité est fermée, donc la condition nécessaire est acquise. La condition suffisante résulte du fait que la diagonale $\Delta \subset G \times G$ est l'image de l'unité par l'application continue $G \times G \to G$, $(a, b) \mapsto ab^{-1}$. □

Corollaire. *Tout groupe de Lie est un groupe topologique séparé.* □

En fait, la différentiabilité de l'application (2) est superflue dans la définition des groupes différentiables.

Proposition 1. *Si pour un groupe G muni d'une structure de variété différentiable, l'application* (1) *est différentiable, l'application* (2) *l'est aussi, et, par suite, le groupe G est un groupe de Lie.*

A noter que cette proposition n'est pas valable pour les groupes topologiques.

La clef de la démonstration de la proposition 1 est donnée par le lemme suivant emprunté à la théorie des variétés différentiables:

Lemme 2. *Soient M, N et R des variétés différentiables et*

$$\varphi: M \times R \to N$$

une application différentiable telle que pour tout point $r \in R$ l'application

$$\varphi_r: M \to N, \quad x \mapsto \varphi(x, r), \quad x \in M,$$

est un difféomorphisme de M sur N. Alors l'application

$$\psi: N \times R \to M,$$

définie par la formule

$$\psi(y, r) = \varphi_r^{-1}(y), \quad y \in N, \quad r \in R,$$

est une application différentiable.

Démonstration. Supposons que les applications

$$\Phi: M \times R \to N \times R \quad \text{et} \quad \Psi: N \times R \to M \times R$$

sont définies respectivement par les formules

$$\Phi(x, r) = (\varphi(x, r), r) = (\varphi_r(x), r), \quad x \in M, \quad r \in R,$$
$$\Psi(y, r) = (\psi(y, r), r) = (\varphi_r^{-1}(y), r), \quad y \in N, \quad r \in R.$$

Il est clair que ces applications sont différentiables si et seulement si les applications φ et ψ le sont respectivement. Ainsi Φ est différentiable par hypothèse. Prouvons que Ψ est différentiable.

Remarquons à cet effet que par définition

$$(\Psi \circ \Phi)(x, r) = \Psi(\varphi_r(x), r) = (\varphi_r^{-1}(\varphi_r(x)), r) = (x, r)$$

pour tout point $(x, r) \in M \times R$ et de façon analogue

$$(\Phi \circ \Psi)(y, r) = \Phi(\varphi_r^{-1}(y), r) = (\varphi_r(\varphi_r^{-1}(y)), r) = (y, r)$$

pour tout point $(y, r) \in N \times R$. Cela signifie que les applications Φ et Ψ sont réciproques l'une de l'autre et par suite bijectives.

Donc, affirmer que Ψ est différentiable revient à affirmer que l'application différentiable bijective Φ est un difféomorphisme.

Mais il est évident qu'une application différentiable bijective est un difféomorphisme si et seulement si elle est un difféomorphisme local, c'est-à-dire est une application étale (sa différentielle est un isomorphisme en tout point).

Tout se ramène donc au calcul en chaque point $(a, r) \in M \times R$ de la différentielle $(d\Phi)_{(a, r)}$ de l'application Φ qui sera traitée comme une application linéaire de la forme *)

$$\mathsf{T}_a(M) \oplus \mathsf{T}_r(R) \to \mathsf{T}_b(N) \oplus \mathsf{T}_r(R), \quad \text{où} \quad b = \varphi(a, r). \tag{3}$$

Chaque application (3) est donnée par une matrice de la forme

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{4}$$

où

$$A: \mathsf{T}_a(M) \to \mathsf{T}_b(N), \quad B: \mathsf{T}_r(R) \to \mathsf{T}_b(N),$$
$$C: \mathsf{T}_a(M) \to \mathsf{T}_r(R), \quad D: \mathsf{T}_r(R) \to \mathsf{T}_r(R)$$

---

*) Par $\mathsf{T}_x(M)$ on désigne l'espace tangent à la variété M en son point x.

sont des applications linéaires qui se construisent de manière évidente. En particulier, si (3) est l'application $(d\Phi)_{(a,\,r)}$, alors l'application A n'est autre que la différentielle en a de l'application $\varphi_r: M \to N$, l'application C, la différentielle d'une application constante (donc C est une application nulle), l'application D, la différentielle de l'application identique (donc D est une application identique aussi). Donc, la matrice (4) associée à la différentielle $(d\Phi)_{(a,\,r)}$ est de la forme

$$\begin{pmatrix} (d\varphi_r)_a & B \\ 0 & \mathrm{id} \end{pmatrix}$$

(l'application B nous importe peu). La différentielle $(d\varphi_r)_a$ étant un isomorphisme par hypothèse, il en est de même de la différentielle $(d\Phi)_{(a,\,r)}$. $\square$

Tout élément a d'un groupe G définit à l'aide des formules

$$L_a x = ax, \quad R_a x = xa, \quad x \in G,$$

deux applications

$$L_a: G \to G \quad \text{et} \quad R_a: G \to G,$$

appelées respectivement *translation à gauche* et *translation à droite de vecteur a*.

Les propriétés suivantes des translations sont évidentes:

$$L_e = R_e = \mathrm{id}, \text{ où } e \text{ est l'unité du groupe } G$$

(ceci exprime que e est l'unité);

$$L_b \circ L_a = L_{ba}, \quad R_b \circ R_a = R_{ab}, \quad L_a \circ R_b = R_b \circ L_a$$

(chacune de ces égalités équivaut à l'associativité de la multiplication dans G).

En particulier, on voit (puisque $L_a \circ L_{a^{-1}} = L_{a^{-1}} \circ L_a = L_e = \mathrm{id}$ et $R_a \circ R_{a^{-1}} = R_{a^{-1}} \circ R_a = R_e = \mathrm{id}$) que chaque translation est une application bijective et de plus que

$$L_a^{-1} = L_{a^{-1}}, \quad R_a^{-1} = R_{a^{-1}}$$

pour tout élément $a \in G$.

Si G est un groupe topologique (resp. différentiable), les applications L_a et R_a sont continues (resp. différentiables), et par suite, sont des homéomorphismes (resp. difféomorphismes).

Nous disposons maintenant de tous les éléments pour prouver la proposition 1.

Démonstration de la proposition 1. La différentiabilité de l'application (1) entraîne celle des translations L_a, et par suite, le fait qu'elles sont des difféomorphismes. L'application $L: (x, a) \mapsto L_a(x) = ax$ n'est autre que l'application (1), donc elle est différentiable. On est ainsi placé dans les conditions du

lemme 1 (pour $M = N = R = G$), et, par suite, en vertu de ce dernier, l'application

$$L' : G \times G \to G,$$

définie par la formule

$$L'(x, a) = L_a^{-1}(x) = a^{-1}x,$$

est différentiable. Pour achever la démonstration il reste à remarquer que l'application $a \mapsto a^{-1}$ est la composée de l'application différentiable $G \to G \times G$, $a \mapsto (e, a)$ et de l'application L'. Donc, elle est aussi différentiable. □

Exemples de groupes de Lie.

Exemple 1. Tout groupe (topologique discret) abstrait est un groupe de Lie pour la structure différentiable qui en fait une variété de dimension zéro.

Exemple 2. Tout espace vectoriel de dimension finie est un groupe de Lie pour l'addition.

Exemple 3. Le cercle unitaire $\mathbb{S}^1$: $|z| = 1$ dont les points sont les nombres complexes $z = e^{i\theta}$ est un groupe de Lie pour la multiplication.

De façon analogue, dans l'espace des quaternions la *sphère unité* $\mathbb{S}^3$ constituée des quaternions ζ tels que $|\zeta| = 1$ est un groupe de Lie pour la multiplication.

On démontre que si une sphère S^n est un groupe de Lie, on a nécessairement $n = 1$ ou $n = 3$, de sorte que $\mathbb{S}^1$ et $\mathbb{S}^3$ sont les seules sphères à admettre une structure de groupe de Lie.

Exemple 4. Le produit direct $G \times H$ de deux groupes différentiables (resp. topologiques) G et H est un groupe différentiable (resp. topologique).

En particulier, le tore T^n, $n \geqslant 1$, est un groupe de Lie.

Exemple 5. Le groupe linéaire complet GL (n) est un groupe de Lie. Le groupe Aut $\mathcal{V}$ des automorphismes (opérateurs linéaires non dégénérés) d'un espace vectoriel $\mathcal{V}$ de dimension n, groupe qui est isomorphe à GL (n), est aussi un groupe de Lie.

Pour obtenir des exemples plus consistants, il nous faut considérer préalablement une construction générale.

Définition 2. On dit qu'une matrice A d'ordre n est *non singulière* si $\det(E + A) \neq 0$. La matrice $A^{\#}$ (lire A dièse):

$$A^{\#} = (E - A)(E + A)^{-1}$$

est dite *Cayley-image de A*.

Il est clair que l'ensemble $\mathfrak{R}(n)^0$ des matrices non singulières est ouvert dans la variété $\mathfrak{R}(n) = \mathbb{R}(n, n)$ des matrices carrées d'ordre $n \times n$ et, par suite, est une variété différentiable.

Proposition 2. *L'application $A \to A^{\#}$ est un autodifféomorphisme involutif de la variété $\mathfrak{R}(n)^0$, c'est-à-dire que pour toute matrice non singulière A la matrice $A^{\#}$ est aussi non singulière, l'application $A \to A^{\#}$ de la variété $\mathfrak{R}(n)^0$ dans elle-même est différentiable et la Cayley-image de la matrice $A^{\#}$ est la matrice A :*

$$A^{\#\#} = A.$$

Démonstration. Soit $B = A^{\#}$. Alors

$$\begin{aligned} E + B &= E + (E - A)(E + A)^{-1} = \\ &= [(E + A) + (E - A)](E + A)^{-1} = \\ &= 2(E + A)^{-1}, \end{aligned}$$

et de façon analogue

$$E - B = 2A(E + A)^{-1}.$$

Donc, *primo*, $\det(E + B) \neq 0$ et, *secundo*,

$$B^{\#} = (E - B)(E + B)^{-1} = A.$$

Il est évident que l'application $A \to A^{\#}$ est différentiable. □

Autres exemples de groupes de Lie.

Exemple 6. Soit, comme toujours, $\mathrm{O}(n)$ le groupe des matrices orthogonales d'ordre n. Montrons que $\mathrm{O}(n)$ peut être naturellement muni d'une structure différentiable et qu'il sera un groupe de Lie pour cette structure.

Soient A une matrice orthogonale non singulière et $B = A^{\#}$. Alors $A^{\top} = A^{-1}$ et par suite

$$\begin{aligned} B^{\top} &= (E + A^{\top})^{-1}(E - A^{\top}) = (E + A^{-1})^{-1}(E - A^{-1}) = \\ &= (E + A^{-1})^{-1} A^{-1} A (E - A^{-1}) = \\ &= (A(E + A^{-1}))^{-1}(A - E) = \\ &= (A + E)^{-1}(A - E) = -(E - A)(E + A)^{-1} = -B. \end{aligned}$$

Inversement, si $B^{\top} = -B$, alors

$$\begin{aligned} A^{\top} &= (E + B^{\top})^{-1}(E - B^{\top}) = (E - B)^{-1}(E + B) = \\ &= (E + B)(E - B)^{-1} = A^{-1}, \end{aligned}$$

et par suite, la matrice A est orthogonale. On voit donc qu'*une condition nécessaire et suffisante pour qu'une matrice non singulière soit orthogonale est que sa Cayley-image soit une matrice antisymétrique.*

Etant donné que les matrices antisymétriques forment un espace vectoriel $\left(\text{de dimension } \frac{n(n-1)}{2}\right)$, l'application $A \mapsto A^{\#}$ peut être traitée comme une application de carte; le voisinage de coor-

données correspondant sera l'ensemble $O(n)^0$ des matrices non singulières orthogonales (qui contient visiblement la matrice unité E), et son image, l'ensemble des matrices non singulières antisymétriques.

Soit maintenant C une matrice orthogonale arbitraire. L'ensemble $O(n)^0 C$ composé des matrices AC, où $A \in O(n)^0$, contient la matrice C dont il est le voisinage. L'application $AC \mapsto A^{\#}$ est une application de carte de ce voisinage sur l'ensemble ouvert des matrices antisymétriques non singulières d'ordre n. Donc, le groupe $O(n)$ est entièrement recouvert de cartes de la forme $O(n)^0 C$. Si $A_1C_1 = A_2C_2$, où $A_1, A_2 \in O(n)^0$, et C_1, C_2 sont des matrices fixes, alors $A_2^{\#} = f(A_1^{\#})$, où f est une fonction matricielle rationnelle dépendant de C_1 et de C_2. La fonction f s'explicite facilement, mais nous nous passerons de cette expression nous contentant simplement d'observer que tout élément de la matrice $A^{\#}$ est une fonction rationnelle, donc différentiable, des éléments de la matrice $A_1^{\#}$. De cette observation il s'ensuit que les cartes de la forme $O(n)^0C$ constituent un atlas car deux quelconques d'entre elles sont cohérentes. La Cayley-image du produit de deux matrices étant visiblement une fonction rationnelle des Cayley-images des facteurs, la structure différentiable sur $O(n)$ est compatible avec la multiplication, c'est-à-dire que *le groupe* $O(n)$ *muni de cette structure différentiable est un groupe de Lie.*

L'artifice utilisé pour les Cayley-images est assez général. En effet, dans tout ce qui précède, les propriétés des matrices orthogonales n'ont été utilisées que pour établir que les Cayley-images des matrices orthogonales non singulières forment un ensemble ouvert d'un espace vectoriel. Donc, *un groupe de matrices est un groupe de Lie si les Cayley-images de ses matrices non singulières forment un ensemble ouvert d'un espace vectoriel de matrices.*

S'agissant des groupes de matrices jouissant de cette propriété, on dira qu'ils *admettent la construction de Cayley*; quant à l'espace vectoriel de matrices correspondant, on l'appellera *Cayley-image du groupe.*

Exemple 7. Une matrice

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$$

d'ordre $n = 2m$ pair est dite *symplectique* si:

a) les matrices $A_1^{\top}A_3$ et $A_2^{\top}A_4$ sont symétriques;

b) l'on a l'égalité

$$A_1^{\top}A_4 - A_3^{\top}A_2 = E.$$

Ces conditions équivalent à l'égalité matricielle

$$A^{\top}JA = J, \tag{5}$$

où

$$J = \begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix}. \tag{6}$$

Dire qu'une matrice est symplectique revient à dire qu'elle respecte la forme antisymétrique bilinéaire

$$(x_1y_{n+1} - x_{n+1}y_1) + \ldots + (x_ny_{2n} - x_{2n}y_n).$$

De façon plus générale, on peut considérer des matrices A vérifiant la relation (5) pour J arbitraire (mais fixe). Si $J = E$, on obtient des matrices orthogonales ($A^{\top}A = E$) et les matrices A satisfaisant (5) sont alors appelées *matrices J-orthogonales.* Ainsi, les matrices symplectiques sont des matrices J-orthogonales correspondant à une matrice J de la forme (6).

De la relation

$$(AB)^{\top} J (AB) = B^{\top} (A^{\top}JA) B$$

il s'ensuit aussitôt que le produit de deux matrices J-orthogonales est une matrice J-orthogonale. Si la matrice J n'est pas dégénérée, en passant aux déterminants dans l'égalité (5), on trouve aussitôt que $\det A = \pm 1$ et, notamment, que toute matrice J-orthogonale est inversible. Sous ces conditions, la matrice inverse A^{-1} est aussi J-orthogonale, puisque

$$(A^{-1})^{\top} JA^{-1} = (A^{-1})^{\top} (A^{\top}JA) A^{-1} = J.$$

Ceci prouve que *si la matrice J n'est pas dégénérée, les matrices J-orthogonales forment un groupe.* En particulier, les matrices symplectiques forment un groupe.

On désignera le groupe des matrices J-orthogonales d'ordre n par $O_J(n)$, et le groupe des matrices symplectiques d'ordre $n = 2m$, par $\mathrm{Sp}\,(m;\ \mathbb{R})$.

Le groupe $\mathrm{Sp}\,(m;\ \mathbb{R})$ s'appelle *groupe symplectique linéaire réel.*

Il est immédiat de voir qu'*une condition nécessaire et suffisante pour qu'une matrice non singulière A soit J-orthogonale est que sa Cayley-image $A^{\#}$ soit J-antisymétrique*, c'est-à-dire que

$$(A^{\#})^{\top} J = -JA^{\#}. \tag{7}$$

En effet, si la relation (5) a lieu, alors

$$\begin{aligned}(A^{\#})^{\top}J &= (E + A^{\top})^{-1} (E - A^{\top}) J = \\ &= (E + JA^{-1}J^{-1})^{-1} (E - JA^{-1}J^{-1}) J = \\ &= J (A^{-1}A + A^{-1})^{-1} (A^{-1}A - A^{-1}) = \\ &= J (A + E)^{-1} (A - E) = -JA^{\#}.\end{aligned}$$

Réciproquement, de (7) il s'ensuit que (on admet que $B = A^{\#}$ et on se sert de la relation $B^{\#} = A$)

$$\begin{aligned} A^{\mathsf{T}} J A &= (E + B^{\mathsf{T}})^{-1} (E - B^{\mathsf{T}}) J \cdot (E - B) (E + B)^{-1} = \\ &= (E + B^{\mathsf{T}})^{-1} \cdot J (E + B) \cdot (E + B)^{-1} (E - B) = \\ &= (E + B^{\mathsf{T}})^{-1} \cdot J (E - B) = (E + B^{\mathsf{T}})^{-1} \cdot (E + B^{\mathsf{T}}) J = \\ &= J. \ \square \end{aligned}$$

Etant linéaire, la condition (7) définit un sous-espace vectoriel dans l'espace des matrices. Ceci prouve que *chaque groupe* $O_J(n)$ (et notamment le groupe Sp $(m; \mathbb{R})$) *admet une construction de Cayley et, par suite, est un groupe de Lie.*

Comme l'espace vectoriel des matrices défini par la condition (7) (avec une matrice J définie par la formule (6)) est de toute évidence de dimension $m(2m + 1)$, il vient, en particulier, que *le groupe* Sp $(m; \mathbb{R})$ *est de dimension* $m(2m + 1)$.

Exemple 8. L'intersection Sp $(m; \mathbb{R}) \cap O(2m)$ s'appelle *groupe symplectique orthogonal.* Les Cayley-images des matrices non singulières de ce groupe sont de la forme

$$\begin{pmatrix} C & D \\ -D & C \end{pmatrix}, \tag{8}$$

où D est une matrice symétrique et C une matrice antisymétrique. Les matrices de la forme (8) formant un espace vectoriel, *le groupe* Sp $(m; \mathbb{R}) \cap O(2m)$ *est un groupe de Lie.* Sa dimension est m^2.

Exemple 9. On peut construire aussi des groupes de Lie avec des matrices complexes. Pour tout $n \geqslant 1$, l'espace vectoriel complexe $\mathbb{C}^n$ peut être identifié à l'espace $\mathbb{R}^{2n}$ en ordonnant (toujours dans le même ordre) les parties réelles et imaginaires des composantes des vecteurs de $\mathbb{C}^n$. Ceci donne à l'espace $\mathbb{C}^n$ (donc à l'un quelconque de ses sous-espaces ouverts) une structure de variété différentiable de dimension $2n$ (qui ne dépend pas bien sûr de l'ordre des parties réelles et imaginaires des composantes des vecteurs de $\mathbb{C}^n$).

L'ensemble $\mathbb{C}(n, m)$ des $n \times m$-matrices complexes s'identifie à $\mathbb{C}^{nm}$, donc est aussi une variété différentiable. Le sous-ensemble GL $(n; \mathbb{C})$ des matrices non dégénérées de $\mathbb{C}(n) = \mathbb{C}(n, n)$ sera aussi une variété différentiable. La variété GL $(n; \mathbb{C})$ est un groupe de Lie (de dimension $2n^2$), puisque les parties réelles et imaginaires des éléments du produit de deux matrices complexes sont des fonctions différentiables des parties réelles et imaginaires des éléments des facteurs.

La notion de Cayley-image et les propriétés de celle-ci se généralisent immédiatement au cas complexe. Ceci vaut également pour les matrices J-orthogonales. On obtient ainsi des *matrices orthogonales complexes* et des *matrices symplectiques complexes.* Ces matrices for-

ment respectivement les groupes de Lie O $(n;\ \mathbb{C})$ et Sp $(m;\ \mathbb{C})$ de dimensions $n(n-1)$ et $2m(2m+1)$.

Exemple 10. Les *matrices A J-unitaires* d'ordre n caractérisées par la relation

$$\bar{A}^{\top}JA = J$$

sont un type différent de matrices complexes. Elles constituent le groupe $U_J(n)$. Pour $J = E$, on obtient des *matrices unitaires* ordinaires et leur groupe U(n). Si J est une matrice (6) (et $n = 2m$), le groupe $U_J(n)$ ne possède pas de notation spéciale. On le désignera par Up (m).

En reproduisant les mêmes raisonnements que précédemment, on établit sans peine qu'*une condition nécessaire et suffisante pour qu'une matrice complexe non singulière A soit J-unitaire est que sa Cayley-image* $A^{\#}$ *vérifie la relation*

$$(\bar{A}^{\#})^{\top} J = -JA^{\#}. \ \square$$

Cette relation étant linéaire (sur le corps $\mathbb{R}$), *le groupe* $U_J(n)$ (*et notamment le groupe* U (n) *et le groupe* Up (m)) *est un groupe de Lie.* □

Le groupe U (n) est de dimension n^2, le groupe Up (m), de dimension $4m^2$.

A noter que *le groupe* U (n) *est canoniquement isomorphe au groupe symplectique orthogonal* Sp $(n;\ \mathbb{R}) \cap$ O $(2n)$. L'isomorphisme est réalisé par la correspondance

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \rightarrow A + iB.$$

L'intersection Sp $(m;\ \mathbb{R}) \cap$ U $(2m)$ est manifestement un groupe symplectique orthogonal:

$$\mathrm{Sp}\,(m;\ \mathbb{R}) \cap \mathrm{U}\,(2m) = \mathrm{Sp}\,(m;\ \mathbb{R}) \cap \mathrm{O}\,(2m)$$

(donc est isomorphe au groupe U (m)).

Exemple 11. L'intersection Sp $(m;\ \mathbb{C}) \cap$ U $(2m)$ s'appelle *groupe symplectique unitaire* (ou simplement *groupe symplectique*) et se note Sp (m). C'est un groupe de Lie de dimension $m(2m+1)$.

L'intersection Sp $(m) \cap$ O $(2m)$ est un groupe symplectique orthogonal.

Exemple 12. Le groupe U (m) peut être interprété comme le groupe des applications linéaires inversibles de l'espace $\mathbb{C}^n$, préservant la forme hermitienne $x_1\bar{y}_1 + \ldots + x_n\bar{y}_n$. De façon analogue, en remplaçant le corps $\mathbb{C}$ par le corps des quaternions $\mathbf{H}$, on peut introduire le groupe $U^{\mathbf{H}}(n)$ des applications linéaires (disons pour la multiplication à gauche) inversibles de l'espace des quaternions

$\mathbb{H}^n$, préservant la forme hermitienne

$$(\xi, \eta) = \xi_1\bar{\eta}_1 + \ldots + \xi_n\bar{\eta}_n,$$

où

$$\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{H}^n, \qquad \eta = (\eta_1, \ldots, \eta_n) \in \mathbb{H}^n.$$

Le groupe $U^{\mathbb{H}}(n)$ peut être naturellement interprété comme un groupe de matrices complexes, puisque tout quaternion ξ peut être identifié à un couple (u, v) de nombres complexes (par la formule $\xi = u + vj$), donc l'espace $\mathbb{H}^n$, à l'espace $\mathbb{C}^{2n}$. Si

$$\xi_1 = x_1 + x_{n+1}j, \ \ldots, \ \xi_n = x_n + x_{2n}j,$$
$$\eta_1 = y_1 + y_{n+1}j, \ \ldots, \ \eta_n = y_n + y_{2n}j,$$

alors (puisque $\overline{u + vj} = \bar{u} - vj$ et $vj = j\bar{v}$)

$$\xi_1\bar{\eta}_1 + \ldots + \xi_n\bar{\eta}_n =$$
$$= [x_1\bar{y}_1 + \ldots + x_n\bar{y}_n + x_{n+1}\bar{y}_{n+1} + \ldots + x_{2n}\bar{y}_{2n}] +$$
$$+ [(x_{n+1}y_1 - x_1y_{n+1}) + \ldots + (x_{2n}y_n - x_ny_{2n})]\, j.$$

Donc, chaque élément du groupe $U^{\mathbb{H}}(n)$ traité comme une matrice complexe préserve la forme hermitienne $x_1\bar{y}_1 + \ldots + x_{2n}\bar{y}_{2n}$ (est une matrice unitaire) et la forme antisymétrique $(x_{n+1}y_1 - x_1y_{n+1}) + \ldots + (x_{2n}y_n - x_ny_{2n})$ (est une matrice symplectique), c'est-à-dire est un élément du groupe symplectique unitaire Sp (n). Réciproquement, si une matrice A est unitaire et symplectique, alors traitée comme une transformation de $\mathbb{H}^n$, elle préserve visiblement la forme $\xi_1\bar{\eta}_1 + \ldots + \xi_n\bar{\eta}_n$. Cette transformation est linéaire. En effet, d'une somme elle donne une somme, de sorte qu'il faut établir seulement sa permutabilité pour la multiplication par un quaternion arbitraire ζ. Or, pour tous vecteurs $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{H}^n$ et $\eta = (\eta_1, \ldots, \eta_n) \in \mathbb{H}^n$, on a

$$(A(\zeta\xi) - \zeta A\xi, A\eta) = (A(\zeta\xi), A\eta) - \zeta(A\xi, A\eta) =$$
$$= (\zeta\xi, \eta) - \zeta(\xi, \eta) = 0,$$

de là il s'ensuit (puisque tout vecteur de $\mathbb{H}^n$ peut être mis sous la forme $A\eta$) que $A(\zeta\xi) = \zeta A\xi$.

Ceci prouve que *le groupe* $U^{\mathbb{H}}(n)$ *est isomorphe au groupe symplectique unitaire* Sp (n). □

On rappelle qu'un espace topologique X est *connexe* s'il ne peut être représenté par la réunion de deux ensembles fermés (ouverts) disjoints non vides, et *connexe par arcs*, s'il existe un *chemin* liant deux points quelconques $a, b \in X$, c'est-à-dire une application continue $u: [0, 1] \to X$ telle que $u(0) = a$, $u(1) = b$. Il est évident (ce qu'on prouve aisément dans tout cours de théorie des nombres

réels) que *le segment* $[0, 1]$ *est connexe*, d'où il s'ensuit aussitôt que *tout espace connexe par arcs est connexe.*

Il est manifeste que l'ensemble de toutes les parties connexes (par arcs) d'un espace topologique X est inductif (vérifie les conditions du lemme de Zorn). Donc, chaque point $a \in X$ appartient à un sous-ensemble connexe maximal C_a, appelé *composante connexe (par arcs)* de l'espace X. On voit sans peine que toute composante connexe $C_a \subset X$ est fermée (mais généralement pas ouverte) dans X. Pour que l'espace X soit connexe (par arcs), il faut et il suffit que $C_a = X$ pour tout point $a \in X$.

Un espace topologique X est *localement connexe* (*par arcs*) si chaque point $a \in X$ possède un système fondamental de voisinages connexes (par arcs), en d'autres termes, si chaque voisinage du point a contient un voisinage connexe (par arcs). Toute variété est un espace localement connexe par arcs.

Dans un espace localement connexe (par arcs) toute composante connexe (par arcs) est visiblement ouverte (car elle contient tout point avec l'un de ses voisinages). De là il s'ensuit, en particulier, que *si un espace est connexe et localement connexe par arcs, il est connexe par arcs.* En d'autres termes, pour les espaces localement connexes par arcs (et notamment pour les variétés) il y a équivalence entre connexité et connexité par arcs.

Exemple 13. Les variétés $\mathbb{R}^n$ et $\mathbb{R}(n, m)$ sont visiblement connexes. La variété $\mathrm{GL}(n)$ ne l'est pas: une matrice de déterminant strictement négatif ne peut être continûment déformée (reliée par un chemin) en une matrice unité. Soit $\mathrm{GL}^+(n)$ l'espace des matrices carrées d'ordre n de déterminant strictement positif. Montrons que *cet espace est connexe.*

Le théorème de décomposition polaire (cf. II, 21) nous dit que tout opérateur non dégénéré est le produit d'un opérateur strictement positif $\boldsymbol{P}$ et d'un opérateur isométrique (orthogonal) $\boldsymbol{U}$. Dans le langage des matrices, cela signifie que toute matrice non dégénérée A est de la forme $A = PU$, où P est la matrice de l'opérateur $\boldsymbol{P}$ et U, la matrice orthogonale de $\boldsymbol{U}$. D'autre part, en vertu du théorème de réduction aux axes principaux, la matrice P est de la forme VDV^{-1}, où V est une matrice orthogonale, D, une matrice diagonale à éléments diagonaux strictement positifs. Donc, $A = VDV^{-1}U$, c'est-à-dire que $A = VDW$, où $W = V^{-1}U$. Mais il est évident que la correspondance $t \to (1 - t) D + tE$ est un chemin continu dans $\mathrm{GL}(n)$ reliant la matrice D à la matrice unité E. En multipliant ce chemin à droite et à gauche par les matrices orthogonales V et W, on obtient un chemin continu reliant dans $\mathrm{GL}(n)$ la matrice A à la matrice orthogonale $B = VW$. Donc, *toute matrice non dégénérée A peut être reliée par un chemin continu dans* $\mathrm{GL}(n)$ *à une matrice orthogonale B.* Si $\det A > 0$, alors $\det B > 0$, c'est-à-dire que $\det B = 1$ (une matrice orthogonale à déterminant strictement

positif est unimodulaire). Donc, pour prouver la connexité du groupe $\mathrm{GL}^+(n)$ il suffit de montrer que *toute matrice orthogonale unimodulaire B peut être reliée par un chemin continu (dans le groupe* $\mathrm{GL}^+(n)$) *à la matrice unité E*. Plus, on prouvera que *cette liaison peut être réalisée dans le groupe* $\mathrm{SO}(n)$ *des matrices orthogonales unimodulaires*. On observera à cet effet que le théorème fondamental des opérateurs orthogonaux (cf. II, 21) nous apprend que tout opérateur orthogonal (une rotation) unimodulaire (c'est-à-dire préservant l'orientation) est la somme directe de l'opérateur identique et de « rotations planes » à matrices de la forme

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

En remplaçant dans chacune de ces matrices l'angle θ par $t\theta$, on obtient une famille continue (un chemin) d'opérateurs orthogonaux reliant cet opérateur (obtenu pour $t = 1$) à l'opérateur identique (obtenu pour $t = 0$). Pour achever la démonstration il reste à passer des opérateurs aux matrices. □

La proposition établie exprime que *le groupe* $\mathrm{GL}(n)$ *est constitué de deux composantes*: le sous-groupe $\mathrm{GL}^+(n)$ et la classe $\mathrm{GL}^-(n)$ suivant $\mathrm{GL}^+(n)$, composée des matrices à déterminants strictement négatifs.

On désignera par G_e la composante de l'unité e de chaque groupe topologique G. Si $a \in G_e$, alors $a \in L_a(G_e)$ (car $e \in G_e$) et, par suite, $G_e \cap L_a(G_e) \neq \varnothing$. G_e et donc $L_a(G_e)$ étant le plus grand ensemble connexe, il vient $L_a(G_e) = G_e$. On démontre de façon analogue que $R_a(G_e) = G_e$ si $a \in G_e$, et que $G_e^{-1} = G_e$. Ceci exprime que G_e est un sous-groupe du groupe G. Bien plus, tout endomorphisme T du groupe G associe à G_e un sous-groupe connexe $T(G_e)$ coupant G_e. Donc, pour les mêmes raisons, $T(G_e) \subset G_e$. Ce qui veut dire que *la composante G_e de l'unité est un sous-groupe complètement invariant du groupe G et notamment est invariante*.

En vertu de ce qui a été dit plus haut, la composante de l'unité du groupe $\mathrm{GL}(n)$ est le groupe $\mathrm{GL}^+(n)$.

A noter que la composante G_e de l'unité de tout groupe différentiable G est automatiquement un groupe différentiable.

Il est naturel de munir le groupe quotient G/G_e de la topologie de l'identification, c'est-à-dire d'une topologie pour laquelle un sous-ensemble $C \subset G/G_e$ est ouvert (resp. fermé) si et seulement si est ouverte (resp. fermée) son image réciproque dans G. Etant donné que l'image réciproque de l'unité du groupe G/G_e est la composante G_e, on voit, en particulier, que l'unité du groupe quotient G/G_e est isolée (est un ensemble ouvert et fermé) ou, en d'autres termes, est discrète si et seulement si la composante G_e est ouverte (fermée, elle

l'est toujours). Ceci aura lieu, en particulier, si le groupe G est localement connexe (par exemple est un groupe différentiable).

Donc, *tout groupe G localement connexe (en particulier, tout groupe différentiable) est l'extension d'un groupe connexe (de la composante G_e de son unité) par le groupe discret G/G_e.*

De ce point de vue, la théorie des groupes localement connexes se ramène à la théorie des groupes connexes et à la théorie des groupes discrets (abstraits).

Pour cette raison, nous admettrons toujours dans la théorie générale que tous les groupes de Lie considérés sont connexes.

Exemple 14. D'après ce qui précède, *le groupe* $\mathrm{SO}(n)$ *est connexe* et, par suite, est la composante de l'unité du groupe $\mathrm{O}(n)$. On voit, en particulier, que *le groupe* $\mathrm{SO}(n)$ *est un groupe de Lie.*

On constate aussi que *le groupe* $\mathrm{O}(n)$ *n'est pas connexe* et qu'il admet deux composantes: le groupe $\mathrm{SO}(n) = \mathrm{O}^{+}(n)$ des matrices (unimodulaires) orthogonales propres et sa classe $\mathrm{O}^{-}(n)$ suivant $\mathrm{O}^{+}(n)$, composée des matrices orthogonales impropres (de déterminant égal à —1).

Exemple 15. A l'inverse, *le groupe* $\mathrm{U}(n)$ *est connexe.* En effet, on sait (cf. II, 21) que tout opérateur unitaire est orthogonalement diagonalisable, et, de plus, toutes ses valeurs propres sont égales à l'unité en module. Dans le langage des matrices, cela signifie que toute matrice unitaire est de la forme UDU^{-1}, où U est une matrice unitaire, D, une matrice diagonale d'éléments $e^{i\theta_k}$. En remplaçant les angles θ_k par $t\theta_k$, on obtient une famille continue (un chemin) de matrices unitaires reliant la matrice considérée (obtenue pour $t = 1$) à la matrice unité (obtenue pour $t = 0$). (Voir plus haut le même raisonnement pour les matrices orthogonales.) □

On peut établir la connexité du groupe $\mathrm{U}(n)$ d'une autre façon, en se servant du lemme général suivant:

Lemme 3. *Un groupe topologique G est connexe s'il contient un sous-groupe connexe H tel que l'espace quotient G/H soit connexe.*

D é m o n s t r a t i o n. Observons tout d'abord que *l'application canonique* $\pi: G \to G/H$ *est ouverte*, c'est-à-dire qu'elle associe un ouvert à un ouvert. En effet, si $U \subset G$, alors par définition de la topologie quotient, l'ensemble $\pi(U) \subset G/H$ est ouvert si et seulement si l'ensemble $\pi^{-1}(\pi(U)) \subset G$ l'est. Mais il est évident que le dernier ensemble est la réunion $\bigcup_{x \in U} xH$ de toutes les classes à gauche xH, $x \in U$, donc il est confondu avec la réunion $\bigcup_{h \in H} Uh$ de toutes les translations de l'ensemble U de vecteur $h \in H$. Par conséquent, si U, donc tout Uh, est ouvert, alors il en est de même de l'ensemble $\pi^{-1}(\pi(U))$, donc de $\pi(U)$.

Supposons maintenant que $G = U \cup V$, où U et V sont des ensembles ouverts non vides. Alors $G/H = \pi(U) \cup \pi(V)$, où les ensembles $\pi(U)$ et $\pi(V)$ sont aussi ouverts et non vides. Donc, $\pi(U) \cap \pi(V) \neq \neq \varnothing$ (car l'espace G/H est connexe par hypothèse). Supposons que $\pi(a) \in \pi(U) \cap \pi(V)$. L'appartenance $\pi(a) \in \pi(U)$ signifie que la classe $\pi(a) = aH$ coupe U, et l'appartenance $\pi(a) \in \pi(V)$, que cette classe coupe V. Ceci étant, $aH = U_1 \cap V_1$, où $U_1 = aH \cap U$ et $V_1 = aH \cap V$ sont ouverts dans aH (et non vides comme on vient de le prouver). Etant donné que aH est connexe (avec H), ceci n'est possible que si $U_1 \cap V_1 \neq \varnothing$, et, par suite, $U \cap V \neq \varnothing$. Donc, le groupe G est connexe. □

Pour appliquer ce lemme, considérons l'application $\mathrm{U}(n) \to \mathbb{C}^n$ qui à chaque matrice associe sa dernière colonne. L'image du groupe $\mathrm{U}(n)$ par cette application est constituée de tous les vecteurs de $\mathbb{C}^n$ de longueur 1, donc elle peut être identifiée à la sphère unité S^{2n-1} de dimension $2n - 1$ de l'espace $\mathbb{R}^{2n} = \mathbb{C}^n$. L'image réciproque de chacun de ces vecteurs dans $\mathrm{U}(n)$ est visiblement une classe suivant le sous-groupe $\mathrm{U}(n-1)$ qui est l'image réciproque du vecteur $(0, 0, \ldots, 0, 1)$. Donc, l'application envisagée induit une application bijective $\mathrm{U}(n)/\mathrm{U}(n-1) \to S^{2n-1}$ de l'espace quotient $\mathrm{U}(n)/\mathrm{U}(n-1)$ sur la sphère S^{2n-1}, qui est (par une vérification automatique) un homéomorphisme (toute application continue bijective sur un compact est un homéomorphisme).

La sphère $S^{2n-1} = \mathrm{U}(n)/\mathrm{U}(n-1)$ étant visiblement connexe, il s'ensuit immédiatement du lemme 2 que le groupe $\mathrm{U}(n)$ est connexe si le groupe $\mathrm{U}(n-1)$ l'est. Le groupe $\mathrm{U}(1)$ s'identifiant canoniquement au groupe $\mathbb{S}^1$, il est connexe, ce qui prouve de nouveau par récurrence la connexité de tous les groupes $\mathrm{U}(n)$.

Exemple 16. On peut appliquer le même raisonnement au groupe symplectique $\mathrm{Sp}(n) = \mathrm{U}^{\mathbb{H}}(n)$. Dans ce cas, l'espace quotient $\mathrm{Sp}(n)/\mathrm{Sp}(n-1)$ s'identifie canoniquement à la sphère unité S^{4n-1} de l'espace $\mathbb{R}^{4n} = \mathbb{H}^n$, donc il est connexe aussi. Le groupe $\mathrm{Sp}(1) = \mathrm{U}^{\mathbb{H}}(1)$ s'identifie au groupe $\mathbb{S}^3$ des quaternions ξ tels que $|\xi| = 1$. Donc, *le groupe* $\mathrm{Sp}(n) = \mathrm{U}^{\mathbb{H}}(n)$ *est connexe pour tout* $n \geqslant 1$.

On peut appliquer le même raisonnement au groupe $\mathrm{SO}(n)$ dont la connexité a été établie plus haut par un autre procédé. En effet, dans ce cas l'espace quotient $\mathrm{SO}(n)/\mathrm{SO}(n-1)$ s'identifie par le même procédé à la sphère unité S^{n-1} de l'espace $\mathbb{R}^n$, sphère qui est connexe pour $n \geqslant 2$. De plus, le groupe $\mathrm{SO}(2)$ des rotations du plan est isomorphe au groupe $\mathbb{S}^1$, donc il est connexe aussi. Ce qui prouve que *le groupe* $\mathrm{SO}(n)$ *est connexe pour* $n \geqslant 2$.

L'analogue du groupe $\mathrm{SO}(n)$ dans le groupe $\mathrm{U}(n)$ est le *sous-groupe* $\mathrm{SU}(n)$ *des matrices unitaires unimodulaires*. Comme, ce qui est immédiat, $\mathrm{SU}(n)/\mathrm{SU}(n-1) = \mathrm{U}(n)/\mathrm{U}(n-1)$ et que le

groupe SU (1) est connexe, car unitaire, en reproduisant le même raisonnement on s'assure que *le groupe* SU (n) *est connexe pour tout* $n \geqslant 1$. Mais nous ne pourrons dire si c'est un groupe de Lie qu'à la leçon 3 après avoir développé les instruments nécessaires.

Idem pour le groupe SL (n) des matrices unimodulaires.

A noter que le groupe SU (n) n'a pas d'analogue dans le groupe U^{H} (n).

LEÇON 2

Champs de vecteurs invariants à gauche.— Parallélisabilité des groupes de Lie.— Courbes intégrales de champs de vecteurs invariants à gauche et sous-groupes à un paramètre.— Foncteur de Lie.— Exemple: groupe d'éléments inversibles d'une algèbre associative.— Fonctions à valeurs dans une algèbre associative.— Sous-groupes à un paramètre du groupe $G(\mathcal{A})$.

On rappelle que par $\mathsf{T}_a(M)$ on désigne l'espace tangent à une variété M en son point a. Les espaces tangents $\mathsf{T}_a(M)$, où a parcourt la variété M, forment une variété différentiable $\mathsf{T}(M)$ de dimension $2n$ (où $n = \dim M$) qui se projette canoniquement sur M. La projection

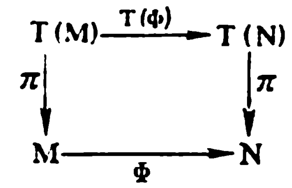
$$\pi: \mathsf{T}(M) \to M$$

associe à tout vecteur A son « point d'application », c'est-à-dire un point $a \in M$ tel que $A \in \mathsf{T}_a(M)$, de sorte que $\mathsf{T}_a(M) = \pi^{-1}(a)$. Les sections de cette projection, c'est-à-dire les applications différentiables

$$X: M \to \mathsf{T}(M), \quad a \mapsto X_a, \quad a \in M,$$

telles que $\pi \circ X = \mathrm{id}$, i.e. $X_a \in \mathsf{T}_a(M)$, s'appellent *champs de vecteurs sur* M. Ces champs de vecteurs engendrent canoniquement un espace vectoriel (de dimension infinie) qui sera désigné par $\mathfrak{a}(M)$.

Les différentielles $(d\Phi)_a: \mathsf{T}_a(M) \to \mathsf{T}_{\Phi a}(N)$ d'une application différentiable arbitraire $\Phi: M \to N$ forment une application différentiable $\mathsf{T}(\Phi): \mathsf{T}(M) \to \mathsf{T}(N)$ qui est justiciable du diagramme commutatif



et les correspondances $M \mapsto \mathsf{T}(M)$, $\Phi \mapsto \mathsf{T}(\Phi)$ sont visiblement un foncteur de la catégorie DIFF des variétés différentiables dans elle-même.

Si Φ est un difféomorphisme, tout champ de vecteurs X de $\mathfrak{a}(M)$ induit le champ

$$\Phi_* X = \mathsf{T}(\Phi) \circ X \circ \Phi^{-1}$$

de $\mathfrak{a}(N)$, et tout champ de vecteurs Y de $\mathfrak{a}(N)$, le champ

$$\Phi^* Y = \mathsf{T}(\Phi)^{-1} \circ Y \circ \Phi$$

de $\mathfrak{a}(M)$. Il est clair que les applications Φ_* et Φ^* sont linéaires et qu'elles sont des isomorphismes réciproques l'un de l'autre d'espaces vectoriels, puisque

$$\Phi_* = (\Phi^*)^{-1} = (\Phi^{-1})^* \quad \text{et} \quad \Phi^* = (\Phi_*)^{-1} = (\Phi^{-1})_*.$$

Si, en particulier, $M = N = G$, où G est un groupe de Lie, pour tout élément $a \in G$ et tout champ de vecteurs $X \in \mathfrak{a}(G)$ sera défini un champ de vecteurs $L_a^* X \in \mathfrak{a}(G)$.

Définition 1. On dit qu'un champ de vecteurs $X \in \mathfrak{a}(G)$ est *invariant à gauche* si $L_a^* X = X$ pour tout élément $a \in G$, c'est-à-dire si

$$(1) \qquad X_b = (dL_{a^{-1}})_{ab}\,(X_{ab}), \quad \forall a, b \in G.$$

Il est évident que les champs de vecteurs invariants à gauche forment un sous-espace de l'espace $\mathfrak{a}(G)$ des champs de vecteurs. On désignera ce sous-espace par $\mathfrak{l}(G)$ ou $\mathfrak{g}$.

Il est immédiat de voir qu'*une condition nécessaire et suffisante pour qu'un champ de vecteurs $X \in \mathfrak{a}(G)$ soit invariant à gauche est que*

$$(2) \qquad X_a = (dL_a)_e\, X_e$$

pour tout élément $a \in G$. En effet, la relation (2) est un cas particulier de la formule (1) (pour $b = e$), et, par suite, la condition nécessaire est remplie si le champ X est invariant à gauche. Réciproquement, si (2) est réalisée, pour tous éléments a, $b \in G$, on a

$$X_{ab} = (dL_{ab})_e\,(X_e) = ((dL_a)_b \circ (dL_b)_e)\,(X_e) = (dL_a)_b\,(X_b),$$

ce qui équivaut à (1). $\square$

De là il s'ensuit qu'*une application linéaire $X \mapsto X_e$ de l'espace $\mathfrak{g}$ dans l'espace tangent $\mathsf{T}_e(G)$ est un isomorphisme.* En effet, pour tout vecteur $A \in \mathsf{T}_e(G)$, l'application $a \mapsto (dL_a)_e A$, $a \in G$, est, comme il est aisé de le voir, un champ de vecteurs sur G (la vérification n'implique que la différentiabilité, laquelle résulte immédiatement de l'expression de cette application en coordonnées locales) jouissant de la propriété (1), donc il est invariant à gauche. Pour achever la

démonstration, il reste à remarquer que l'application $\mathsf{T}_e(G) \to \mathfrak{g}$ obtenue est visiblement réciproque de l'application $X \mapsto X_e$. □

En principe, on identifiera l'espace $\mathfrak{g} = \mathfrak{l}(G)$ à l'espace $\mathsf{T}_e(G)$ par l'application $X \mapsto X_e$.

Comme $\dim \mathsf{T}_e(G) = n$, où $n = \dim G$, on voit, en particulier, que *pour tout groupe de Lie G, l'espace $\mathfrak{g} = \mathfrak{l}(G)$ des champs de vecteurs invariants à gauche est de dimension finie égale à $n = \dim G$.*

Soit $\mathcal{F}(M)$ l'algèbre des fonctions différentiables sur une variété différentiable M. Pour toute fonction $f \in \mathcal{F}(M)$ et tout champ de vecteurs $X \in \mathfrak{a}(M)$, la formule

$$(fX)_a = f(a)\, X_a, \quad a \in M,$$

définit de toute évidence un champ de vecteurs $fX \in \mathfrak{a}(M)$, et, comme le montre une vérification immédiate, l'espace vectoriel $\mathfrak{a}(M)$ est un module sur l'algèbre $\mathcal{F}(M)$ pour l'opération $(f, X) \mapsto fX$. Si ce module est un module libre de rang n, c'est-à-dire si sur M il existe un système $X_1, \ldots, X_n$ de champs de vecteurs (base du $\mathcal{F}(M)$-module $\mathfrak{a}(M)$) tel que tout champ de vecteurs $X \in \mathfrak{a}(M)$ se représente de façon unique sous la forme

$$X = f^1X_1 + \ldots + f^nX_n,$$

où $f^1, \ldots, f^n \in \mathcal{F}(M)$, alors la variété M est dite *parallélisable*.

Proposition 1. *Tout groupe de Lie G est parallélisable.*

D é m o n s t r a t i o n. Plus, nous prouverons que *toute base $X_1, \ldots, X_n$ de l'espace vectoriel $\mathfrak{l}(G)$ est une base du $\mathcal{F}(G)$-module $\mathfrak{a}(G)$.*

Pour tout point $a \in G$, les vecteurs $(X_1)_a, \ldots, (X_n)_a$ forment une base de l'espace vectoriel $\mathsf{T}_a(G)$. Donc, tout vecteur X_a d'un champ de vecteurs arbitraire $X \in \mathfrak{a}(G)$ se décompose de façon unique suivant les vecteurs $(X_1)_a, \ldots, (X_n)_a$. Cela signifie que pour tout champ de vecteurs $X \in \mathfrak{a}(G)$, il existe des fonctions $f^i : a \mapsto f^i(a)$, $a \in G$, telles que $X = f^1X_1 + \ldots + f^nX_n$. Il nous faut donc prouver seulement que $f^k \in \mathcal{F}(G)$ pour tous les $k = 1, \ldots, n$.

Soit $(U, x^1, \ldots, x^n)$ une carte de la variété G. Les champs $X_1, \ldots, X_n$ étant différentiables, il existe sur U des fonctions différentiables $X_1^i, \ldots, X_n^i$, $i = 1, \ldots, n$, telles que

$$X_j = X_j^i \frac{\partial}{\partial x^i} \quad \text{pour tout } j = 1, \ldots, n.$$

Vu, sous ces conditions, que pour tout point $a \in U$, les vecteurs $(X_1)_a, \ldots, (X_n)_a$ forment une base de l'espace $\mathsf{T}_a(G)$, il vient

$$\det (X_j^i) \neq 0 \text{ sur } U,$$

et, par suite, il existe sur U des fonctions différentiables Y_i^k telles que

$$X_j^i Y_i^k = \delta_j^k, \quad i, j, k = 1, \ldots, n.$$

Par hypothèse $X = f^j X_j$, et, par suite,

$$X = f^j X_j^i \frac{\partial}{\partial x^i},$$

autrement dit les fonctions $f^j X_j^i$, $i = 1, \ldots, n$, sont les composantes du champ de vecteurs X dans le système de coordonnées locales $x^1, \ldots, x^n$, donc elles sont différentiables. Mais

$$f^k = f^j \delta_j^k = (f^j X_j^i)\, Y_i^k,$$

et comme les fonctions $f^j X_j^i$ et Y_i^k sont différentiables, il s'ensuit que les fonctions f^k le sont aussi (sur U).

Etant différentiables sur tout voisinage U, les fonctions f^k le sont sur la variété G tout entière. □

On rappelle qu'une courbe différentiable $t \mapsto \varphi(t)$ sur une variété M est une *courbe intégrale* d'un champ de vecteurs si

$$\frac{d\varphi(t)}{dt} = X_{\varphi(t)} \quad \forall t.$$

Une courbe intégrale est dite *maximale* si elle n'est pas la restriction d'une courbe définie sur un plus long intervalle de l'axe réel. Du classique théorème d'existence et d'unicité de la solution d'un système d'équations différentielles à seconds membres différentiables et des propriétés élémentaires des espaces séparés, il s'ensuit immédiatement que si une variété M est séparée (ce qui, on le sait, est automatique pour un groupe de Lie), alors *pour tout point $a \in M$, il existe une courbe intégrale maximale φ_a du champ X passant pour $t = 0$ par le point a, c'est-à-dire telle que $\varphi_a(0) = a$.*

Le champ de vecteurs X est dit *complet* si pour tout point $a \in M$, la courbe φ_a est définie sur l'axe $\mathbb{R}$ tout entier.

Il est immédiat qu'*une condition nécessaire et suffisante pour qu'un champ de vecteurs X sur un groupe de Lie G soit invariant à gauche est que pour deux points quelconques $a, b \in G$ l'on ait*

$$\varphi_{ab} = L_a \circ \varphi_b, \tag{3}$$

c'est-à-dire que $\varphi_{ab}(t) = a\varphi_b(t)$, $t \in \mathbb{R}$. En effet, pour tout $a \in G$ fixe, la formule $\psi_b(t) = \varphi_{a^{-1}b}(t)$ définit pour tout point $b \in G$ une courbe $t \mapsto \psi_b(t)$ qui passe par b pour $t = 0$ et il est clair qu'en posant

$$Y_b = \left.\frac{d\psi_b(t)}{dt}\right|_{t=0},$$

on obtient un champ de vecteurs $Y\colon b \mapsto Y_b$ sur G. Ceci étant, les règles de calcul des vecteurs tangents à des courbes différentiables

en tout point $b \in G$ nous donnent

$$Y_b = \frac{d\psi_b(t)}{dt}\bigg|_{t=0} = \frac{d\,(L_a \circ \varphi_{a^{-1}b}(t))}{dt}\bigg|_{t=0} =$$

$$= (dL_a)_{a^{-1}b}\left(\frac{d\varphi_{a^{-1}b}(t)}{dt}\bigg|_{t=0}\right) = (dL_a)_{a^{-1}b}(X_{a^{-1}b}).$$

Si (3) est réalisée, donc $\psi_b = \varphi_b$ (et, par suite, $Y_b = X_b$), alors $X_b = (dL_a)_{a^{-1}b}(X_{a^{-1}b})$ et, en particulier, $X_a = (dL_a)_e(X_e)$. Donc, le champ X est invariant à gauche. Réciproquement, si le champ X est invariant à gauche (donc vérifie la relation (1)), $Y_b = X_b$ pour tout point $b \in G$, c'est-à-dire que $Y = X$. Or, il est clair que les courbes $t \mapsto \psi_b(t)$ sont des courbes intégrales du champ Y (qui sont automatiquement maximales), donc, en vertu de l'égalité $Y = X$, elles sont confondues avec les courbes intégrales $t \mapsto \varphi_b(t)$ du champ X. Par conséquent, $\varphi_b(t) = a\varphi_{a^{-1}b}(t)$, ce qui équivaut à (3). □

Définition 2. Une courbe différentiable $\beta : \mathbb{R} \to G$ est un *sous-groupe à un paramètre* d'un groupe de Lie G si

$$\beta(t+s) = \beta(t)\,\beta(s)$$

pour tous $t, s \in \mathbb{R}$. En d'autres termes, un sous-groupe à un paramètre est un homomorphisme du groupe additif $\mathbb{R}$ des réels (traité comme un groupe de Lie) dans le groupe de Lie G.

On observera qu'un sous-groupe à un paramètre n'est pas un sous-ensemble mais une application.

Il est évident que tout sous-groupe à un paramètre β passe par l'unité e du groupe G pour $t = 0$:

$$\beta(0) = e.$$

Proposition 2. *Tout sous-groupe à un paramètre β est une courbe intégrale d'un champ de vecteurs X invariant à gauche.*

Démonstration. La formule

$$\varphi_a(t) = a\beta(t), \quad a \in G, \quad t \in \mathbb{R},$$

définit sur G une courbe différentiable $t \mapsto \varphi_a(t)$ passant par le point a pour $t = 0$. Posons

$$X_a = \frac{d\varphi_a(t)}{dt}\bigg|_{t=0}.$$

Une vérification immédiate montre que l'application $a \mapsto X_a$ est différentiable, c'est-à-dire est un champ de vecteurs sur G et que les courbes φ_a sont les courbes intégrales de ce champ. En particulier, $\varphi_e = \beta$ est une courbe intégrale. Enfin, le champ X est invariant à gauche, puisque

$$\varphi_{ab}(t) = (ab)\,\beta(t) = a\,(b\beta(t)) = a\varphi_b(t). \ \square$$

La réciproque de la proposition 2 est également vraie.

Proposition 3. *La courbe intégrale maximale β d'un champ de vecteurs $X \in \mathfrak{g}$ invariant à gauche, qui passe par le point e pour $t = 0$, est un sous-groupe à un paramètre d'un groupe de Lie G (et, en particulier, est définie sur $\mathbb{R}$ tout entier).*

Démonstration. Le champ X étant invariant à gauche, ses courbes intégrales φ_a satisfont la relation (3). Donc, en particulier, l'intervalle I_a de $\mathbb{R}$ sur lequel est définie la courbe intégrale φ_a est confondu avec l'intervalle $I = I_e$ sur lequel est définie la courbe intégrale $\beta = \varphi_e$. D'autre part, étant donné que pour tout $s \in \mathbb{R}$ fixe la courbe $t \mapsto \varphi_e(t + s)$ est visiblement une courbe intégrale du champ X passant par le point $b = \varphi_e(s)$ et donc que $\varphi_e(t + s) = \varphi_b(t)$, il vient

$$(4) \quad \beta(s + t) = \beta(t + s) = \varphi_e(t + s) = \\ = \varphi_b(t) = b\varphi_e(t) = \varphi_e(s)\,\varphi_e(t) = \beta(s)\,\beta(t)$$

pour tous $s, t \in I$ tels que $s + t \in I$. Pour prouver la proposition 3, il nous faut donc montrer seulement que la courbe β est définie sur $\mathbb{R}$ tout entier, c'est-à-dire que $I = \mathbb{R}$.

Supposons que $I \neq \mathbb{R}$. Pour tout $t \in \mathbb{R}$, il existe un entier n tel que $\frac{t}{n} \in I$. Prolongeons la courbe β à la droite $\mathbb{R}$ tout entière en posant

$$\beta(t) = \beta\left(\frac{t}{n}\right)^n \quad \text{si } \frac{t}{n} \in I.$$

Cette définition est correcte. En effet, si $\frac{t}{n} \in I$ et $\frac{t}{m} \in I$, alors $\frac{t}{nm} \in I$, et, par suite, en vertu de (4),

$$\beta\left(\frac{t}{n}\right)^n = \left[\beta\left(\frac{t}{nm}\right)^m\right]^n = \left[\beta\left(\frac{t}{nm}\right)^n\right]^m = \beta\left(\frac{t}{m}\right)^m.$$

Donc, la courbe β est prolongée à la droite $\mathbb{R}$ tout entière. Il est évident que cette courbe est différentiable et vérifie la relation (4) pour tous $t, s \in \mathbb{R}$, c'est-à-dire est un sous-groupe à un paramètre. Si l'on prouve que la courbe β est une courbe intégrale du champ X sur la droite $\mathbb{R}$ tout entière, on sera conduit à une contradiction avec l'hypothèse $I \neq \mathbb{R}$.

Soient $t_0 \in \mathbb{R}$ et $a = \beta(t_0)$. L'action du vecteur tangent $\frac{d\beta(t_0)}{dt}$ à la courbe β en a sur la fonction *) $f \in \Theta_a(G)$ est définie par la

*) M étant une variété différentiable et $a \in M$, on désigne par $\Theta_a(M)$ l'ensemble des fonctions *différentiables en a*, c'est-à-dire définies et différentiables dans un voisinage (dépendant de la fonction considérée) du point a. (A proprement parler, par $\Theta_a(M)$ on devrait entendre l'espace vectoriel des germes, mais ne soyons pas trop pédants.)

Par $\Theta(M)$ on désignera l'ensemble des fonctions f définies et différentiables dans un ouvert $W(f) \subset M$ (dépendant de la fonction envisagée).

formule

$$\frac{d\beta(t_0)}{dt} f = \frac{d(f\circ\beta)(t)}{dt}\Big|_{t=t_0}.$$

De façon analogue, le vecteur tangent $\frac{d\beta(0)}{dt}$ en e agit sur la fonction $f \in \mathcal{O}_e(G)$ à l'aide de la formule

$$\frac{d\beta(0)}{dt} f = \frac{d(f\circ\beta)(t)}{dt}\Big|_{t=0}.$$

Donc, pour toute fonction $f \in \mathcal{O}_a(G)$, on a

$$\left[(dL_a)_e \frac{d\beta(0)}{dt}\right] f = \frac{d\beta(0)}{dt}(f\circ L_a) =$$

$$= \frac{d(f\circ L_a\circ\beta)(t)}{dt}\Big|_{t=0} = \frac{df(a\beta(t))}{dt}\Big|_{t=0} =$$

$$= \frac{df(\beta(t+t_0))}{dt}\Big|_{t=0} = \frac{df(\beta(t))}{dt}\Big|_{t=t_0} = \frac{d\beta(t_0)}{dt} f,$$

c'est-à-dire que

$$(dL_a)_e \frac{d\beta(0)}{dt} = \frac{d\beta(t_0)}{dt}.$$

Mais pour $t \in I$ la courbe β est une courbe intégrale du champ X. En particulier,

$$\frac{d\beta(0)}{dt} = X_{\beta(0)} = X_e.$$

D'autre part, l'invariance à gauche du champ X entraîne $(dL_a)_e X_e = X_a = X_{\varphi(t_0)}$. Donc

$$X_{\varphi(t_0)} = \frac{d\beta(t_0)}{dt},$$

de sorte que la courbe β est bien une courbe intégrale du champ de vecteurs X. □

Corollaire. *Tout champ de vecteurs invariant à gauche est complet.* □

En vertu des propositions 2 et 3, les champs invariants à gauche $X \in \mathfrak{l}(G)$ et les sous-groupes à un paramètre β sont en correspondance biunivoque. Donc, pour éléments de l'espace vectoriel $\mathfrak{l}(G)$, on peut, si l'on veut, prendre des sous-groupes à un paramètre.

En combinant ce qui vient d'être dit on obtient le

Théorème 1. *L'espace* $\mathfrak{g} = \mathfrak{l}(G)$ *admet les trois interprétations équivalentes suivantes:*

I) *L'espace* $\mathfrak{g}$ *est composé des champs de vecteurs* X *invariants à gauche sur le groupe de Lie* G.

II) *L'espace* $\mathfrak{g}$ *est composé des vecteurs tangents* A *au groupe* G *au point* e *(l'unité du groupe* G*).*

III) *L'espace* $\mathfrak{g}$ *est composé des sous-groupes à un paramètre* β *du groupe* G.

L'équivalence entre I) *et* II) *est donnée par la correspondance*

$$X \mapsto X_e,$$

l'équivalence entre III) *et* II), *par la correspondance*

$$\beta \mapsto \frac{d\beta(0)}{dt};$$

et, enfin, l'équivalence entre I) *et* III), *par la correspondance*

$$X \mapsto \varphi_e,$$

où φ_e *est une courbe intégrale du champ* X, *passant par le point* e *pour* $t = 0$. □

La première et la deuxième interprétation définissent les opérations linéaires dans $\mathfrak{g}$ pour lesquelles $\mathfrak{g}$ est un espace vectoriel de dimension n. Nous montrerons dans la leçon 4 comment obtenir ces opérations linéaires pour la troisième interprétation.

Soit $\Phi: G \to H$ un homomorphisme de groupes de Lie. Etant donné que $\Phi(e) = e$, la différentielle $\mathsf{T}_e\Phi = (d\Phi)_e$ de cet homomorphisme au point e est une application linéaire de l'espace $\mathsf{T}_e(G) = \mathfrak{l}(G)$ dans l'espace $\mathsf{T}_e(H) = \mathfrak{l}(H)$.

Il est clair que les correspondances $G \mapsto \mathfrak{l}(G)$, $\Phi \mapsto (d\Phi)_e$ constituent un foncteur

(5) $$\mathfrak{l}: \text{GR-DIFF} \to \text{LIN}_f(\mathbb{R})$$

de la catégorie GR-DIFF des groupes de Lie dans la catégorie $\text{LIN}_f(\mathbb{R})$ des espaces vectoriels de dimension finie sur le corps $\mathbb{R}$.

Définition 3. Le foncteur (5) sera appelé *foncteur de Lie*. Dans un but d'unification des notations, on désignera l'application $(d\Phi)_e$ aussi par le symbole $\mathfrak{l}(\Phi)$.

Nous avons défini l'application $\mathfrak{l}(\Phi)$ en traitant les espaces $\mathfrak{l}(G)$ comme des espaces tangents en e. Une question se pose: comment définir l'application $\mathfrak{l}(\Phi)$ dans les autres interprétations?

On rappelle que des champs de vecteurs $X \in \mathfrak{a}(M)$ et $Y \in \mathfrak{a}(M)$ sont par définition Φ-*liés*, où Φ est une application différentiable de M sur N, si

(6) $$Y_{\Phi a} = (d\Phi)_a(X_a) \qquad \forall a \in M.$$

Si les champs de vecteurs sont interprétés comme des opérateurs différentiels linéaires $f \mapsto Xf$ (de différentiations sur M), la Φ-liaison signifie que toute fonction f (définie et différentiable sur un ouvert de la variété N) satisfait la relation $X(f \circ \Phi) = Yf \circ \Phi$.

Si $\Phi: M \to N$ est un difféomorphisme, alors chaque champ $Y \in \mathfrak{a}(N)$ est Φ-lié au champ $\Phi^*Y \in \mathfrak{a}(M)$.

Proposition 4. *Si les éléments des espaces* $\mathfrak{l}(G)$ *sont traités comme des sous-groupes à un paramètre, l'application* $\mathfrak{l}(\Phi)$ *est donnée par la formule*

$$\mathfrak{l}(\Phi)(\beta) = \Phi \circ \beta.$$

Pour tout champ de vecteurs invariant à gauche $X \in \mathfrak{l}(G)$ *et tout homomorphisme* $\Phi\colon G \to H$, *il existe sur le groupe* H *un seul champ de vecteurs* Y *invariant à gauche qui est* Φ*-lié au champ* X. *Si les espaces* $\mathfrak{l}(G)$ *sont interprétés comme des espaces de champs de vecteurs invariants à gauche, l'image* $\mathfrak{l}(\Phi)X$ *du champ* $X \in \mathfrak{l}(G)$ *par l'application* $\mathfrak{l}(\Phi)$ *est précisément le champ* Y.

D é m o n s t r a t i o n. Par définition

$$(d\Phi)_e \left(\frac{d\beta(t)}{dt}\bigg|_{t=0} \right) = \frac{d(\Phi \circ \beta)(t)}{dt}\bigg|_{t=0}.$$

Donc, si l'on identifie le sous-groupe à un paramètre β au vecteur $A = \frac{d(\beta(t))}{dt}\Big|_{t=0}$, le sous-groupe à un paramètre $\Phi \circ \beta$ s'identifie au vecteur $(d\Phi)_e A = \mathfrak{l}(\Phi)A$. Ceci prouve la première proposition.

La deuxième proposition résulte immédiatement du fait évident que pour les champs de vecteurs invariants à gauche $X \in \mathfrak{l}(G)$ et $Y \in \mathfrak{l}(H)$, la relation (6) (pour $M = G$ et $N = H$) équivaut à l'égalité

$$Y_e = (d\Phi_e)\, X_e. \ \square$$

Il est clair que les espaces $\mathsf{T}_e(G)$ et $\mathsf{T}_e(G_e)$ sont confondus. Cela veut dire que

$$\mathfrak{l}(G) = \mathfrak{l}(G_e).$$

Donc, en étudiant un foncteur de Lie, on peut se limiter aux seuls groupes différentiables connexes G.

Illustrons les notions introduites sur un exemple concret important.

Rappelons préalablement quelques notions classiques d'algèbre qui seront constamment utilisées dans la suite.

Soit $\mathbb{K}$ un corps arbitraire (pour nous ce sera le corps $\mathbb{R}$ des réels), et soit $\mathcal{A}$ un espace vectoriel sur $\mathbb{K}$. Supposons qu'à deux éléments $x, y \in \mathcal{A}$ est associé un troisième élément $z \in \mathcal{A}$ noté xy et nommé *produit* des éléments x et y. Alors, chaque élément $a \in \mathcal{A}$ définira deux applications

$$L_a\colon x \mapsto ax, \quad R_a\colon x \mapsto xa$$

de $\mathcal{A}$ dans $\mathcal{A}$.

Définition 4. Un espace vectoriel $\mathcal{A}$ muni de la multiplication $x, y \mapsto xy$ s'appelle *algèbre* (sur le corps $\mathbb{K}$) si pour tout élément

$a \in \mathscr{A}$ les applications $L_a: \mathscr{A} \to \mathscr{A}$ et $R_a: \mathscr{A} \to \mathscr{A}$ sont linéaires, c'est-à-dire si

$$(7) \qquad a(x+y) = ax + ay, \quad (x+y)a = xa + ya$$

pour tous éléments $x, y \in \mathscr{A}$ et

$$(8) \qquad k(ax) = a(kx), \quad k(xa) = (kx)a$$

pour tout élément $x \in \mathscr{A}$ et tout élément $k \in \mathbb{K}$.

La condition (7) (combinée aux quatre premiers axiomes de l'espace vectoriel) signifie que l'espace $\mathscr{A}$ muni de l'addition et de la multiplication est un anneau. On peut donc dire qu'une algèbre est un anneau muni d'une structure d'espace vectoriel dans lequel est remplie la condition (8), c'est-à-dire la condition

$$(9) \qquad k(xy) = (kx)y = x(ky)$$

qui doit être réalisée pour tous éléments $x, y \in \mathscr{A}$ et tout élément $k \in \mathbb{K}$.

On appelle *homomorphisme* d'algèbres une application linéaire d'une algèbre dans une autre associant la multiplication à la multiplication (i.e. est un homomorphisme d'anneaux).

Il est clair que les algèbres (sur un corps $\mathbb{K}$ donné) et leurs homomorphismes forment une catégorie qui sera désignée par ALG.

Un sous-espace vectoriel $\mathscr{B}$ d'une algèbre $\mathscr{A}$ s'appelle *sous-algèbre* de $\mathscr{A}$ si $xy \in \mathscr{B}$ pour tous éléments $x, y \in \mathscr{B}$. Il est clair qu'une sous-algèbre est automatiquement une algèbre.

Définition 5. Une algèbre munie d'une multiplication associative s'appelle *algèbre associative*.

Les algèbres associatives forment une sous-catégorie complète ALG-ASS de la catégorie ALG. (Une sous-catégorie $\mathbf{B}$ d'une catégorie $\mathbf{C}$ est dite *complète* si pour tous objets $B_1, B_2 \in \mathbf{B}$, chaque morphisme $B_1 \to B_2$ de $\mathbf{C}$ appartient à $\mathbf{B}$.)

Toute sous-algèbre d'une algèbre associative est associative.

Les algèbres associatives qui admettent une unité (c'est-à-dire un élément e tel que $ae = ea = a$ pour tout élément $a \in \mathscr{A}$) seront appelées *algèbres unitaires*. Elles forment une sous-catégorie complète ALG_0-ASS de la catégorie ALG-ASS.

L'*algèbre des matrices* $\mathbb{K}(n)$ est une algèbre unitaire.

On dit qu'un élément a d'une algèbre unitaire $\mathscr{A}$ est *inversible* s'il existe dans $\mathscr{A}$ un élément a^{-1} tel que $aa^{-1} = a^{-1}a = e$. L'ensemble $G(\mathscr{A})$ des éléments inversibles de l'algèbre $\mathscr{A}$ est visiblement un groupe pour la multiplication.

Il est évident qu'un élément a est inversible si et seulement si l'opérateur linéaire L_a l'est, c'est-à-dire si $\det L_a \neq 0$ dans le cas où l'algèbre $\mathscr{A}$ est de dimension finie.

Pour $\mathbb{K} = \mathbb{R}$ il s'ensuit de là que *si l'algèbre $\mathscr{A}$ est de dimension*

finie, l'ensemble $G(\mathcal{A})$ est ouvert dans $\mathcal{A}$ et, par suite, est une variété différentiable (de dimension $n = \dim \mathcal{A}$).

Donc, $G(\mathcal{A})$ est à la fois un groupe et une variété différentiable. La multiplication sur ce groupe étant bilinéaire, celui-ci est visiblement différentiable et, par suite, *est un groupe de Lie* (cf. proposition 1 de la leçon 1).

Déterminons $\mathfrak{l}(G(\mathcal{A}))$.

Rappelons-nous à cet effet que pour tout espace vectoriel $\mathcal{V}$ de dimension finie (traité comme une variété différentiable) et pour tout point $v \in \mathcal{V}$, l'espace tangent $\mathsf{T}_v(\mathcal{V})$ s'identifie canoniquement à $\mathcal{V}$: l'isomorphisme

$$\mathcal{V} \to \mathsf{T}_v(\mathcal{V}) \tag{10}$$

associe à tout vecteur $a \in \mathcal{V}$ le vecteur tangent à la courbe $t \mapsto v + ta$ en $t = 0$. Donc, on a, en particulier, $\mathsf{T}_e(\mathcal{A}) = \mathcal{A}$. D'autre part, $G(\mathcal{A})$ étant ouvert dans $\mathcal{A}$, il vient $\mathsf{T}_e(G(\mathcal{A})) = \mathsf{T}_e(\mathcal{A})$. Donc (l'espace vectoriel $\mathfrak{l}(G)$ est traité ici comme l'espace tangent $\mathsf{T}_e(G)$)

$$\mathfrak{l}(G(\mathcal{A})) = \mathcal{A}. \tag{11}$$

Plaçons-nous dans le cadre des deux autres interprétations de l'espace $\mathfrak{l}(G)$ et essayons de discuter l'égalité (11).

Soient $\mathcal{V}$ un espace vectoriel de dimension finie et $A : \mathcal{V} \to \mathcal{V}$ un opérateur linéaire. L'opérateur A étant visiblement une application différentiable de la variété différentiable $\mathcal{V}$ dans elle-même, sa différentielle $(dA)_v : \mathsf{T}_v(\mathcal{V}) \to \mathsf{T}_{Av}(\mathcal{V})$ est définie en tout point $v \in \mathcal{V}$. Par définition, cette différentielle associe au vecteur tangent en v à la courbe $t \mapsto v + ta$ le vecteur tangent en Av à la courbe $t \mapsto A(v + ta) = Av + tAa$. Ceci exprime que la différentielle $(dA)_v$ est confondue avec l'opérateur A, puisque $\mathsf{T}_v(\mathcal{V}) = \mathcal{V}$ et $\mathsf{T}_{Av}(\mathcal{V}) = \mathcal{V}$. Donc, *la différentielle d'un opérateur linéaire n'est autre que cet opérateur.*

On voit, en particulier, que $(dL_a)_e = L_a$ pour tout élément $a \in A$.

D'autre part, en vertu des identifications $\mathsf{T}_v(\mathcal{V}) = \mathcal{V}$ et $\mathsf{T}_{Av}(\mathcal{V}) = \mathcal{V}$, tout champ de vecteurs X sur $G(\mathcal{A})$ n'est rien d'autre qu'une application différentiable de $G(\mathcal{A})$ sur $\mathcal{A}$. La condition d'invariance à gauche (1) pour le champ de vecteurs X ainsi traité est de la forme

$$X_b = L_{a^{-1}} X_{ab}, \quad \text{où} \quad a,\ b \in G(\mathcal{A}), \tag{12}$$

d'où il s'ensuit pour $b = e$ que $X_a = L_a X_e$, c'est-à-dire que le champ est de la forme $a \mapsto ab$, où $b = X_e \in \mathcal{A}$. Vu qu'un tel champ vérifie manifestement la condition (12), en modifiant les notations, on trouve que *les champs de vecteurs invariants à gauche $G(\mathcal{A}) \to \mathcal{A}$ sur le groupe de Lie $G(\mathcal{A})$ sont de la forme $x \mapsto xa$, $x \in G(\mathcal{A})$*, où a est un élément arbitraire de l'algèbre $\mathcal{A}$.

La discussion de l'égalité (11) dans le cadre de la troisième interprétation des vecteurs de $\mathfrak{l}(G(\mathcal{A}))$, c'est-à-dire de leur interprétation comme des sous-groupes à un paramètre, implique l'introduction de quelques notions préliminaires.

Une norme définie sur une algèbre $\mathcal{A}$ (sur le corps $\mathbb{R}$) est dite *multiplicative* si

$$\| ab \| \leqslant \| a \| \cdot \| b \|$$

quels que soient les éléments $a, b \in \mathcal{A}$.

Lemme 1. *Toute algèbre $\mathcal{A}$ de dimension finie sur le corps $\mathbb{R}$ est munie d'une norme multiplicative.*

D é m o n s t r a t i o n. Si dans $\mathcal{A}$ est définie une base $e_1, \ldots$ $\ldots, e_n$, alors la formule

$$(13) \qquad \| a \| = \max (| a^1 |, \ldots, | a^n |),$$

où $a^1, \ldots, a^n$ sont les coordonnées de l'élément a dans la base $e_1, \ldots, e_n$, définit manifestement une norme dans $\mathcal{A}$. Nous allons montrer que si la base $e_1, \ldots, e_n$ est convenablement choisie, la norme (13) est multiplicative.

Supposons d'abord que la base $e_1, \ldots, e_n$ est arbitraire et que

$$e_i e_j = c_{ij}^k e_k, \quad i, j, k = 1, \ldots, n.$$

Pour tous éléments $a = a^i e_i$ et $b = b^i e_i$, on a

$$\| ab \| = \| c_{ij}^k a^i b^j e_k \| = \max_k | c_{ij}^k a^i b^j | \leqslant$$

$$\leqslant \sum_{i, j=1}^{n} \max_k | c_{ij}^k | \cdot \max_p | a^p | \cdot \max_q | b^q | = C \cdot \| a \| \cdot \| b \|,$$

où $C = n^2 \max\limits_{i, j, k} | c_{ij}^k |$. Donc, pour la norme

$$\| a \| = \lambda \max (| a^1 |, \ldots, | a^n |),$$

où $\lambda > C$ $\left(\text{cette norme est la norme (13) associée à la base } \frac{1}{\lambda} e_1, \ldots \right.$ $\left. \ldots, \frac{1}{\lambda} e_n \right)$, on a l'inégalité

$$\| ab \| \leqslant \frac{C}{\lambda} \| a \| \cdot \| b \| \leqslant \| a \| \cdot \| b \|,$$

c'est-à-dire que cette norme est multiplicative. □

A noter que dans le lemme 1 on ne suppose pas l'associativité de l'algèbre $\mathcal{A}$.

Nous utiliserons cette remarque dans la suite. Pour l'instant appliquons le lemme 1 à une algèbre $\mathcal{A}$ unitaire (d'unité e) associative

de dimension finie. Introduisons pour tout élément $a \in \mathcal{A}$ la séri

(14) $$e + ta + \frac{t^2a^2}{2} + \ldots + \frac{t^na^n}{n!} + \ldots$$

Cette série *converge absolument* pour toute norme multiplicative, c'est-à-dire que la série

$$\| e \| + \| ta \| + \left\| \frac{t^2a^2}{2} \right\| + \ldots + \left\| \frac{t^na^n}{n!} \right\| + \ldots$$

converge (puisqu'elle est majorée par la série de $e^{t\| a \|}$). Mais la démonstration classique (produite généralement pour les séries à termes scalaires mais valable *in extenso* pour les séries à termes vectoriels) montre *que toute série absolument convergente est convergente* (en norme, donc en coordonnées dans un espace vectoriel de dimension finie). Donc, *la série* (14) *converge.*

La somme de la série (14) est désignée par e^{ta}, et la fonction $t \mapsto e^{ta}$ à valeurs dans $\mathcal{A}$ est dite *fonction exponentielle dans l'algèbre* $\mathcal{A}$. (La lettre e n'a manifestement rien de commun avec l'unité e de l'algèbre $\mathcal{A}$.)

En particulier, pour $\mathcal{A} = \mathbb{R}(n)$, on obtient la *fonction exponentielle matricielle* $t \mapsto e^{tA}$, $A \in \mathbb{R}(n)$.

On peut reproduire pratiquement toutes les constructions de l'analyse élémentaire pour les fonctions $t \mapsto a(t)$ à valeurs dans $\mathcal{A}$. Par exemple, la *dérivée* $t \mapsto a'(t)$ de la fonction $t \mapsto a(t)$ est donnée par la formule

(15) $$a'(t) = \lim_{\Delta t \to 0} \frac{a(t+\Delta t) - a(t)}{\Delta t}$$

et si la fonction $t \mapsto a(t)$ est suffisamment différentiable, alors

(16) $$a(t) = a(t_0) + (t - t_0)\, a'(t_0) + O(t - t_0)^2.$$

Dans le même temps, toute fonction $t \mapsto a(t)$ peut être traitée comme une courbe dans $\mathcal{A}$ munie d'une structure de variété différentiable et l'on peut ainsi envisager le vecteur $\frac{da(t)}{dt}$ qui lui est tangent en t et qui, en vertu de l'isomorphisme (10), peut être regardé comme un vecteur de $\mathcal{A}$.

Il apparaît que ces deux définitions sont équivalentes, c'est-à-dire que

$$a'(t) = \frac{da(t)}{dt} \quad \text{pour tout } t.$$

En effet, en vertu de la formule (15), le vecteur tangent à la courbe $t \mapsto a(t)$ au point $t = t_0$ est confondu avec le vecteur tangent à la courbe $t \mapsto a(t_0) + (t - t_0)\, a'(t_0)$, qui, en vertu de l'isomorphisme (10), s'identifie au vecteur $a'(t_0)$. □

Mais la définition (15) est indiscutablement plus commode en pratique, car elle donne immédiatement toutes les formules usuelles du calcul différentiel (par exemple, la formule de dérivation du produit $(a(t)\, b(t))' = a'(t)\, b(t) + a(t)\, b'(t)$) si seulement sont prises les précautions dues à une éventuelle non-commutativité de la multiplication dans l'algèbre $\mathcal{A}$ (si tel est le cas, la formule de la dérivée de la fonction $t \to a^{-1}(t)$ devient $(a^{-1}(t))' = -a^{-1}(t) \times$ $\times\, a'(t)\, a^{-1}(t))$.

Si les valeurs d'une fonction $t \mapsto a(t)$ à valeurs dans $\mathcal{A}$ commutent, c'est-à-dire si $a(t)\, a(s) = a(s)\, a(t)$ pour tous t et s, alors aucune restriction n'est à faire. Donc, en particulier, pour tout polynôme

$$f(X) = a_0 + a_1 X + \ldots + a_m X^m$$

et toute fonction $a \mapsto a(t)$ à valeurs dans $\mathcal{A}$ dont les valeurs commutent, on a la formule

$$\frac{d}{dt} f(a(t)) = f'(a(t))\, a'(t), \tag{17}$$

où

$$f'(X) = a_1 + 2a_2 + \ldots + m a_m X^{m-1}.$$

Cette formule est valable aussi lorsque $f(X)$ est la somme de la série entière

$$f(X) = a_0 + a_1 X + \ldots + a_n X^n + \ldots, \tag{18}$$

puisque la permutation des deux passages à la limite est visiblement légitime (à condition bien sûr que $\| a(t) \|$ soit situé dans le disque de convergence de la série (18)).

Les séries de fonctions à valeurs dans $\mathcal{A}$ sont justiciables aussi des règles usuelles de dérivation terme à terme. C'est le cas notamment de la série (14). Donc

$$\frac{de^{ta}}{dt} = a + ta^2 + \ldots + \frac{t^{n-1}a^n}{(n-1)!} + \ldots =$$

$$= a\left(e + ta + \ldots + \frac{t^{n-1}a^{n-1}}{(n-1)!} + \ldots\right) = ae^{ta}.$$

Par conséquent, *la fonction exponentielle* $t \mapsto e^{ta}$ *est telle que*

$$\frac{de^{ta}}{dt} = ae^{ta} \tag{19}$$

pour tout t.

De là il s'ensuit que *la solution (à valeurs dans* $\mathcal{A}$*) de l'équation différentielle*

$$\frac{dx(t)}{dt} = ax(t) \tag{20}$$

vérifiant la condition initiale

$$x(0) = c$$

est définie par la formule

$$x(t) = e^{ta}c.$$

En effet, en vertu de la formule (17),

$$x'(t) = (e^{ta})'c = ae^{ta}c = ax(t)$$

et $x(0) = c$. D'autre part, l'équation (20) se ramène à un système d'équations différentielles linéaires à coefficients constants; donc elle admet une solution $x(t)$ et une seule. □

Il est aisé de voir maintenant que *pour tous s et t, on a*

$$e^{(t+s)a} = e^{ta}e^{sa}. \tag{21}$$

En effet, pour tout s fixe, la fonction $t \mapsto x(t) = e^{(t+s)a}$ est solution de l'équation (20) avec la condition initiale $x(0) = e^{sa}$. Donc, $x(t) = e^{ta}e^{sa}$. □

De la relation (21) il s'ensuit, en particulier, que *la fonction $t \mapsto e^{ta}$ est une fonction à valeurs commutables.* Donc (cf. formule (17)) pour toute série entière (18), on a

$$\frac{d}{dt} f(e^{ta}) = f'(e^{ta})\, ae^{ta} \tag{22}$$

(à condition bien sûr que la série de $f(e^{ta})$ converge absolument).

Revenons maintenant à l'espace vectoriel $\mathfrak{l}(G)$ pour $G = G(\mathcal{A})$.

Les sous-groupes à un paramètre du groupe $G(\mathcal{A})$ ne sont rien d'autre que des fonctions $x \mapsto x(t)$ différentiables à valeurs dans $\mathcal{A}$, vérifiant la relation

$$x(s + t) = x(s)\,x(t), \quad s, t \in \mathbb{R}. \tag{23}$$

D'après ce qui précède, on trouve sans aucune peine la solution de cette équation fonctionnelle. En effet, en dérivant la relation (23) par rapport à s et en posant ensuite $s = 0$, on retrouve pour $x(t)$ l'équation différentielle (20) avec $a = x'(0)$. Donc, en tenant compte de la condition initiale $x(0) = e$, on trouve $x(t) = e^{ta}$. Vu qu'en vertu de (21), cette solution vérifie (23), il vient que *tout sous-groupe à un paramètre du groupe $G(A)$ est de la forme $t \mapsto e^{ta}$.*

En désignant le sous-groupe à un paramètre $t \mapsto e^{ta}$ par β_a, on obtient la correspondance biunivoque $a \mapsto \beta_a$ entre les éléments de l'algèbre $\mathcal{A}$ et les sous-groupes à un paramètre du groupe de Lie $G(\mathcal{A})$. Ceci n'est autre que la correspondance (11) de la troisième interprétation.

LEÇON 3

Groupes de Lie des matrices admettant la construction de Cayley.— Généralisation de la construction de Cayley.— — Groupes possédant des ln-images.— Algèbres de Lie.— Exemples d'algèbres de Lie.— Algèbres de Lie des champs de vecteurs.— Algèbre de Lie des groupes de Lie.— Exemple: l'algèbre de Lie du groupe des éléments inversibles d'une algèbre associative.— Groupes de Lie localement isomorphes.— Groupuscules de Lie.— Foncteur de Lie sur la catégorie des groupuscules de Lie.

Les résultats obtenus à la fin de la leçon précédente s'étendent manifestement au groupe linéaire complet GL $(n) = G(\mathfrak{R}(n))$. On voit, en particulier, que *les fonctions exponentielles matricielles* $t \mapsto e^{tA}$ *et elles seules sont des sous-groupes à un paramètre du groupe* GL (n).

Définition 1. On dira qu'un sous-groupe G du groupe GL (n) est un *groupe de Lie de matrices* si:

a) G est muni d'une structure différentiable pour laquelle il est un groupe de Lie;

b) l'injection $\iota: G \to$ GL (n) est différentiable (et, par suite, est un homomorphisme de groupes de Lie).

Tout sous-groupe à un paramètre du groupe G est nécessairement un sous-groupe à un paramètre du groupe GL (n), donc est de la forme $t \mapsto e^{tA}$. Ceci définit une injection $\mathfrak{l}(G) \to \mathfrak{l}(\mathrm{GL}(n)) = \mathfrak{R}(n)$ qui n'est autre (cf. proposition 5 de la leçon 2) que l'application $\mathfrak{l}(\iota)$. Donc, *l'espace vectoriel* $\mathfrak{g} = \mathfrak{l}(G)$ *s'identifie canoniquement à un sous-espace de l'espace vectoriel* $\mathfrak{R}(n)$.

Tout groupe admettant la construction de Cayley (cf. leçon 1), par exemple le groupe $\mathrm{O}_J(n)$ des matrices J-orthogonales, est un groupe de matrices. Par définition, le sous-groupe à un paramètre matriciel $t \mapsto e^{tA}$ est un sous-groupe à un paramètre du groupe $\mathrm{O}_J(n)$ si et seulement si pour tout $t \in \mathfrak{R}$, on a

$$(e^{tA})^{\mathsf{T}} J e^{tA} = J.$$

En dérivant par rapport à t et en posant $t = 0$, on obtient la relation

$$A^{\mathsf{T}}J + JA = 0,$$

qui exprime que la matrice A est une matrice J-antisymétrique. Réciproquement, pour toute matrice A J-antisymétrique, l'application $t \to e^{tA}$ est un sous-groupe à un paramètre du groupe $O_J(n)$.

Pour établir ce fait, on se servira de l'analogue matriciel de la formule classique élémentaire

$$e^a = \lim_{m\to\infty}\left(1 + \frac{a}{m}\right)^m.$$

Montrons (en remplaçant 1 par e) que *cette formule est valable dans toute algèbre associative de dimension finie.* En effet, comme

$$\frac{1}{m^k}\binom{m}{k} = \frac{m(m-1)\ldots(m-k+1)}{\underbrace{m\cdot m\cdot\ldots\cdot m}_{k \text{ facteurs}}}\cdot\frac{1}{k!} \leqslant \frac{1}{k!},$$

il vient pour toute norme multiplicative

$$\left\| e^a - \left(e + \frac{a}{m}\right)^m \right\| = \left\| \sum_{k=0}^{\infty}\left(\frac{1}{k!} - \frac{1}{m^k}\binom{m}{k}\right) a^k \right\| \leqslant$$

$$\leqslant \sum_{k=0}^{\infty}\left(\frac{1}{k!} - \frac{1}{m^k}\binom{m}{k}\right)\|a\|^k =$$

$$= e^{\|a\|} - \left(1 + \frac{\|a\|}{m}\right)^m,$$

donc

$$\lim_{m\to\infty}\left\| e^a - \left(e + \frac{a}{m}\right)^m \right\| = 0,$$

car

$$\left(1 + \frac{\|a\|}{m}\right)^m \to e^{\|a\|}. \ \square$$

On voit que pour tout t

$$Je^{tA} = J\lim_{m\to\infty}\left(E + \frac{tA}{m}\right)^m = \lim_{m\to\infty} J\left(E + \frac{tA}{m}\right)^m =$$

$$= \lim_{m\to\infty}\left(E - \frac{tA^{\mathsf{T}}}{m}\right)^m J = e^{-tA^{\mathsf{T}}}J,$$

car $Jf(A) = f(-A^{\mathsf{T}})J$ pour tout polynôme $f(A)$ de la matrice A. Donc

$$(e^{tA})^{\mathsf{T}}Je^{tA} = (e^{tA})^{\mathsf{T}}e^{-tA^{\mathsf{T}}}J = J,$$

de sorte que $e^{tA} \in O_J(n)$.

Ceci prouve que *pour le groupe* $O_J(n)$, *l'espace vectoriel des matrices J-antisymétriques est le sous-espace* $\mathfrak{l}\,(O_J(n))$ *de l'espace* $\mathbb{R}\,(n)$.

En comparant cette assertion avec le résultat obtenu dans l'exemple 7 de la leçon 1, on constate que *le sous-espace* $\mathfrak{l}\,(O_J(n))$ *est confondu avec la Cayley-image du groupe* $O_J(n)$. Il s'avère que c'est un fait général.

Proposition 1. *Si un groupe de matrices* $G \subset \mathrm{GL}\,(n)$ *admet la construction de Cayley (donc est un groupe de Lie de matrices), l'espace vectoriel* $\mathfrak{l}\,(G)$ *est confondu avec la Cayley-image* $G^{\#}$ *du groupe* G.

D é m o n s t r a t i o n. Supposons que $A \in \mathfrak{l}\,(G)$, c'est-à-dire que l'application $t \mapsto e^{tA}$ est un sous-groupe à un paramètre du groupe G. L'ensemble G^0 des matrices non singulières de G étant un voisinage de l'unité E du groupe G, il existe un $\varepsilon > 0$ tel que pour $|t| < \varepsilon$, la matrice e^{tA} est non singulière et, par suite, est définie sa Cayley-image

$$(e^{tA})^{\#} = (E - e^{tA})\,(E + e^{tA})^{-1} \in G^{\#}.$$

L'espace $G^{\#}$ étant vectoriel, il contient aussi la matrice

$$\left.\frac{d\,(e^{tA})^{\#}}{dt}\right|_{t=0} = \lim_{t \to 0} \frac{(e^{tA})^{\#}}{t}.$$

Mais, d'autre part

$$\frac{d\,(e^{tA})^{\#}}{dt} = -Ae^{tA}\,(E + e^{tA})^{-1} + (E - e^{tA})\,\frac{d\,(E + e^{tA})^{-1}}{dt},$$

donc

$$\left.\frac{d\,(e^{tA})^{\#}}{dt}\right|_{t=0} = -\frac{1}{2}\,A,$$

et, par suite, $A \in G^{\#}$.

Ceci prouve que $\mathfrak{l}\,(G) \subset G^{\#}$, donc que $\mathfrak{l}\,(G) = G^{\#}$, puisque les espaces vectoriels $\mathfrak{l}\,(G)$ et $G^{\#}$ sont de même dimension ($\dim G$). □

D'après les exemples traités dans la leçon 1, il résulte de la proposition 1 que *les éléments de l'espace* $\mathfrak{l}\,(G)$ *sont :*

les matrices antisymétriques d'ordre n pour le groupe orthogonal $O\,(n)$ (*ou, ce qui est équivalent, pour le groupe* $SO\,(n)$) ;

les matrices de la forme

$$\begin{pmatrix} A & B \\ C & -A^{\top} \end{pmatrix},$$

où B et C sont des matrices symétriques d'ordre m et A, une matrice arbitraire, pour le groupe symplectique réel $\mathrm{Sp}\,(m, \mathbb{R})$;

les matrices de la forme

$$\begin{pmatrix} A & -C \\ C & A \end{pmatrix},$$

où A *est antisymétrique et* C, *symétrique, pour le groupe symplectique orthogonal* Sp $(m) \cap$ O $(2m)$;

les matrices antihermitiennes pour le groupe unitaire U (n);

les matrices

$$\begin{pmatrix} A & B \\ C & -\bar{A}^{\mathrm{T}} \end{pmatrix},$$

où B *et* C *sont des matrices hermitiennes d'ordre* m *et* A, *une matric arbitraire, pour le groupe* Up (m);

les matrices

$$\begin{pmatrix} A & B \\ -\bar{B} & \bar{A} \end{pmatrix},$$

où A *est une matrice antihermitienne et* B, *une matrice symétrique d'ordre* m, *pour le groupe symplectique* Sp (m). □

L'affirmation qu'un groupe admettant la construction de Cayley est un groupe de Lie de matrices n'est pas intrinsèquement liée à l'application de Cayley $A \mapsto A^{\#}$ et peut être généralisée.

Comme plus haut, il nous sera plus commode d'identifier l'espace $\mathbb{R}^{n^2}$ à l'espace $\mathbb{R}(n)$ des matrices carrées d'ordre n.

Proposition 2. *Un sous-groupe* G *du groupe* GL (n) *est un groupe de Lie de matrices s'il existe un difféomorphisme* $f: V \to \mathring{V}$ *d'un voisinage* V *de la matrice unité du groupe* GL (n) *sur un ensemble ouvert* $\mathring{V}$ *de l'espace* $\mathbb{R}(n)$, *tel que l'ensemble* $f(G \cap V)$ *soit l'intersection de l'ensemble* $\mathring{V}$ *et d'un sous-espace vectoriel* $G^{\#}$ *de l'espace* $\mathbb{R}(n)$:

$$f(G \cap V) = G^{\#} \cap \mathring{V}.$$

D é m o n s t r a t i o n. Supposons que $m = \dim G^{\#}$ et que $\varphi: G^{\#} \to \mathbb{R}^m$ est un isomorphisme quelconque de $G^{\#}$ sur $\mathbb{R}^m$. Supposons par ailleurs que $U = G \cap V$ et $\mathring{U} = \varphi(G^{\#} \cap \mathring{V})$. Alors l'ensemble $\mathring{U}$ est un ouvert de $\mathbb{R}^m$ et l'application $h = \varphi \circ f$ est une application bijective de U sur $\mathring{U}$. En d'autres termes, le couple (U, h) est une carte sur G.

Supposons maintenant que A est une matrice arbitraire de G et que $U_A = L_A(U)$ et $h_A = h \circ L_A^{-1}$. Le couple (U_A, h_A) est aussi une carte sur G. Comme $A \in U_A$, les ensembles U_A recouvrent G. D'autre part, si $U_A \cap U_B \neq \varnothing$, l'application $h_B \circ h_A^{-1}$ sur

$h_A(U_A \cap U_B)$ sera la restriction du difféomorphisme

$$h \circ L_B^{-1} \circ L_A \circ h^{-1} = \varphi \circ f \circ L_{B^{-1}A} \circ f^{-1} \circ \varphi^{-1},$$

donc sera elle-même un difféomorphisme. Par conséquent, les cartes (U_A, h_A) forment un atlas. Ceci munit G d'une structure différentiable pour laquelle G est visiblement un groupe de Lie de matrices. □

Le cas d'un groupe admettant la construction de Cayley s'obtient lorsque V est l'ensemble des matrices non singulières de G et l'application $f: V \to \mathring{V}$, l'application de Cayley (et, par suite, l'espace vectoriel $G^\#$ est la Cayley-image du groupe G).

La proposition 1 s'étend aussi au cas général traité si l'on exige que le difféomorphisme $f: V \to \mathring{V}$ soit *analytique*, c'est-à-dire si sont remplies les conditions suivantes:

a) il existe un nombre R et une norme matricielle $\| \ \|$ tels que $\| A - E \| < R$ pour toute matrice $A \in V$;

b) il existe une série

$$f(z) = a_0 + a_1(z-1) + \ldots + a_m(z-1)^m + \ldots,$$

convergente pour $|z - 1| < R$, telle que pour toute matrice $A \in V$

$$f(A) = a_0E + a_1(A - E) + \ldots + a_m(A - E)^m + \ldots$$

(cette égalité a un sens en raison de la condition a));

c) le nombre $a_1 = f'(1)$ est non nul.

Proposition 3. *Si pour un sous-groupe G du groupe* GL (n), *il existe un difféomorphisme analytique* $f: V \to \mathring{V}$ *remplissant les conditions de la proposition* 2, *alors l'espace vectoriel* $\mathfrak{l}(G)$ *est confondu avec l'espace vectoriel* $G^\#$ *de la proposition* 2.

Démonstration (comparer avec la démonstration de la proposition 1). Soient $t \mapsto e^{tA}$ un sous-groupe à un paramètre du groupe G, $\varepsilon > 0$ un nombre tel que la matrice e^{tA} appartienne à V pour $|t| < \varepsilon$. Alors $e^{tA} \in G \cap V$, et, par suite, $f(e^{tA}) \in G^\# \cap \mathring{V}$. Donc $\frac{df(e^{tA})}{dt} \in G^\#$ et, en particulier,

$$\left.\frac{df(e^{tA})}{dt}\right|_{t=0} \in G^\#.$$

Or, la formule (22) de la leçon précédente nous dit que

$$\left.\frac{df(e^{tA})}{dt}\right|_{t=0} = f'(e^{tA})Ae^{tA}|_{t=0} = a_1A,$$

puisque

$$f'(z) = a_1 + 2a_2(z-1) + \ldots + ma_m(z-1)^{m-1} + \ldots$$

et donc $f'(E) = a_1E$. Par conséquent, $a_1A \in G^\#$ et $A \in G^\#$, puisque par hypothèse $a_1 \neq 0$.

Ceci prouve que $\mathfrak{l}(G) \subset G^\#$. Donc $\mathfrak{l}(G) = G^\#$, puisque ces espaces sont de même dimension. □

Pour expliciter le difféomorphisme f, considérons la série matricielle

$$\ln A = (A - E) - \frac{1}{2}(A - E)^2 + \ldots + \frac{(-1)^{m+1}}{m}(A - E)^m + \ldots,$$

qui converge pour $\| A - E \| < 1$ (où $\| \quad \|$ est une norme matricielle multiplicative, par exemple la norme $\| A \| = n \cdot \max\limits_{i,j} a_{ij}$). Un calcul trivial répétant un calcul classique sur les séries numériques montre que $e^{\ln A} = A$ pour $\| A - E \| < 1$ (c'est-à-dire lorsque la matrice $\ln A$ est définie).

Il est intéressant de noter au contraire que l'égalité $\ln e^A = A$ peut ne pas être réalisée même lorsque la matrice $\ln e^A$ est définie (en ce sens que la série $\ln B$ converge pour la matrice $B = e^A$). En effet, si

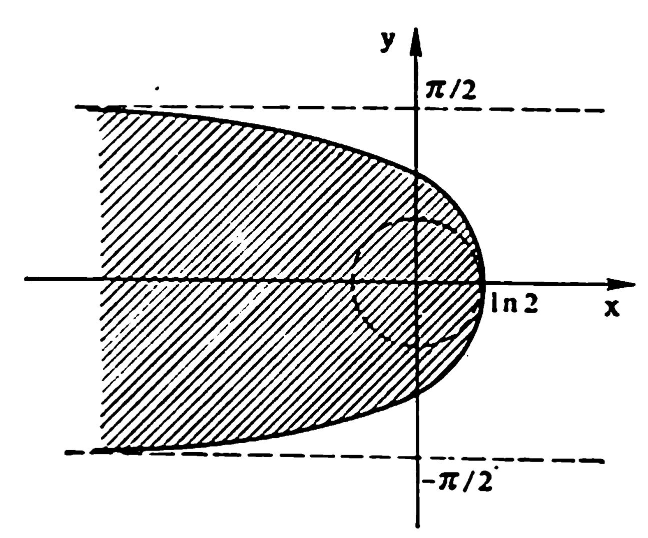
$$A = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix},$$

un calcul immédiat montre que

$$e^A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

et, par suite, $e^A = E$ pour $\theta = 2\pi$. Donc, la matrice $\ln e^A$ est définie et égale à 0 et non pas à A.

Du reste, cette situation se présente aussi pour les nombres complexes. Par exemple, $e^{2\pi i} = 1$ et $\ln e^{2\pi i} = 0$. La cause est connue. La condition $| e^z - 1 | < 1$ définit dans le plan de la variable complexe $z = x + iy$ un système dénombrable de domaines se déduisant l'un de l'autre par des translations de vecteur $2\pi i$, le domaine contenu dans la bande $| y | < \frac{\pi}{2}$ étant borné par la courbe $e^x = 2 \cos y$ (cf. figure). Par ailleurs, la transformation formelle de la



série de $\ln e^z$ en la série de z n'a de sens, de même du reste que toute transformation des séries, que dans le disque de convergence correspondant, qui dans le cas étudié est le plus grand disque de centre $z_0 = 0$ à être contenu dans le domaine envisagé. Comme le rayon de ce disque est égal à $\ln 2$ (ce qui est immédiat), l'égalité $\ln e^z = z$ n'est sûrement réalisée que pour $|z| < \ln 2$.

Il est désormais clair que ces mêmes transformations formelles conviennent aussi pour la série $\ln e^A$ et par suite *l'égalité* $\ln e^A = A$ *a lieu à fortiori pour* $\|A\| < \ln 2$.

On a ainsi prouvé que l'application $\ln: A \to \ln A$ est un difféomorphisme d'un voisinage V de la matrice unité du groupe GL (n) sur un voisinage $\mathring{V}$ de la matrice nulle de l'espace vectoriel $\mathfrak{R}(n)$ (le difféomorphisme réciproque étant $\exp: A \mapsto e^A$).

On dira qu'*un sous-groupe* $G \subset$ GL (n) *admet une* ln-*image* si dans $\mathfrak{R}(n)$ il existe un sous-espace vectoriel $G^\flat$ (lire G bémol) tel que

$$\ln (G \cap V) = G^\flat \cap \mathring{V}.$$

Un tel sous-groupe est un groupe de Lie de matrices en vertu de la proposition 2 et l'espace $G^\flat$ est confondu avec l'espace $\mathfrak{g} = \mathfrak{l}(G)$ en vertu de la proposition 3.

Contrairement à la construction de Cayley, cette construction permet de prouver immédiatement que *les groupes* SL(n) *et* SU(n) *des matrices unimodulaires sont des groupes de Lie de matrices*. En effet, il est classique que $\det e^A = e^{\operatorname{Tr} A}$, où $\operatorname{Tr} A$ est la *trace* de la matrice A (c'est-à-dire la somme de ses éléments diagonaux). Donc, la condition d'unimodularité de la matrice e^A est équivalente à la condition linéaire $\operatorname{Tr} A = 0$. □

(On voit aisément qu'il suffit de démontrer l'égalité $\det e^A = e^{\operatorname{Tr} A}$ uniquement pour les matrices de Jordan ou, à la rigueur, triangulaires. Mais, pour une telle matrice, la matrice e^A est aussi triangulaire et ses éléments diagonaux sont de la forme $e^{a_1}, \ldots, e^{a_n}$, où $a_1, \ldots, a_n$ sont les éléments diagonaux de la matrice A. Donc, $\det e^A = e^{a_1} \ldots e^{a_n} = e^{a_1 + \ldots + a_n} = e^{\operatorname{Tr} A}$.)

L'avantage essentiel de la ln-construction sur la construction de Cayley réside dans son universalité.

Proposition 4. *Tout groupe de Lie de matrices G admet une* ln-*image.*

Démonstration. D'après ce qui a été dit plus haut, le seul prétendant au rôle de $G^\flat$ est l'espace vectoriel $\mathfrak{l}(G)$. Montrons qu'il possède effectivement la qualité requise.

Soient comme plus haut V et $\mathring{V}$ des voisinages respectivement de la matrice unité et de la matrice nulle, tels que la fonction $A \to \ln A$ définisse le difféomorphisme $\ln: V \to \mathring{V}$ de difféomorphisme réci-

proque $\exp: \mathring{V} \to V$. Pour toute matrice $A \in \mathfrak{l}(G) \cap \mathring{V}$ on a alors $e^A \in G \cap V$ (puisque $e^{tA} \in G$ pour tout t). Comme $\ln e^A = A$, il vient que $\mathfrak{l}(G) \cap \mathring{V} \subset \ln (G \cap V)$.

Réciproquement, supposons que $B \in G \cap V$. Alors est définie la matrice $A = \ln B \in \mathring{V}$. Considérons sur GL (n) le champ de vecteurs invariant à gauche $Y: P \mapsto PA$. La restriction $X = Y|_G$ du champ Y à G est manifestement un champ de vecteurs différentiable invariant à gauche sur G (est un élément de l'espace $\mathfrak{l}(G)$) qui est ι-lié au champ Y, où $\iota: G \to \mathrm{GL}(n)$ est une injection. En vertu de la proposition 4 de la leçon 2, cela signifie que $\mathfrak{l}(\iota)X = Y$. Donc, en vertu des identifications générales, le champ X s'identifie à la matrice A. Donc $A \in \mathfrak{l}(G)$. Ce qui prouve que $\ln (G \cap V) \subset \subset \mathfrak{l}(G) \cap \mathring{V}$.

En définitive $\ln (G \cap V) = \mathfrak{l}(G) \cap \mathring{V}$ C.q.f.d. □

Donc, *un groupe de matrices est un groupe de Lie si et seulement s'il admet une* ln-*image*. Le passage à la ln-image pour ainsi dire linéarise le groupe, ce qui a pour effet d'en faciliter l'étude. Etant donné que l'espace $\mathfrak{l}(G)$ (qui est confondu avec sa ln-image pour les groupes de matrices) est défini pour tout groupe de Lie, il est logique d'escompter que le foncteur de Lie $\mathfrak{l}: G \mapsto \mathfrak{l}(G)$ joue en théorie des groupes de Lie un rôle similaire à celui du foncteur ln en théorie des groupes de Lie de matrices. Il s'avère que c'est bien le cas et ce fait est la pierre angulaire de la théorie des groupes de Lie. Notre ouvrage sera consacré essentiellement à l'examen de ces problèmes.

Définition 2. On appelle *algèbre de Lie* une algèbre $\mathfrak{g}$ munie d'une loi de multiplication anticommutative, c'est-à-dire que

$$xy = -yx, \quad \forall x, y \in \mathfrak{g}$$

et telle que pour tout triple (x, y, z) de $\mathfrak{g}$ l'on ait la relation

$$(xy)z + (yz)x + (zx)y = 0, \tag{1}$$

dite *identité de Jacobi*.

Les algèbres de Lie sont généralement désignées par des lettres gothiques minuscules.

Pour tout élément a d'une algèbre de Lie $\mathfrak{g}$, les applications L_a et R_a ne diffèrent que par leur signe ($L_a = -R_a$). L'application L_a est généralement désignée par ad a pour les algèbres de Lie.

Comme les algèbres associatives, les algèbres de Lie forment une sous-catégorie complète de la catégorie ALG.

Nous désignerons cette sous-catégorie par ALG-LIE.

Les algèbres de Lie à dimension finie sur le corps $\mathbb{R}$ occuperont

4—01642

une place privilégiée dans notre exposé. Elles forment une catégorie désignée par ALG_f-LIE.

Observons que toute sous-algèbre d'une algèbre de Lie est une algèbre de Lie.

La construction suivante nous permet d'obtenir un grand nombre d'algèbres de Lie.

Soit $\mathcal{A}$ une algèbre associative arbitraire.

Le *crochet de Lie* de deux éléments $x, y \in \mathcal{A}$ se définit par la formule

$$[x, y] = xy - yx.$$

Il est clair que $[x, y] = -[y, x]$. Par ailleurs,

$$[[x, y], z] + [[y, z], x] + [[z, x], y] =$$
$$= (xy - yx) z - z (xy - yx) + (yz - zy) x - x (yz - zy) + \\ + (zx - xz) y - y (zx - xz) = 0$$

pour tout triple (x, y, z) de $\mathcal{A}$. Ceci exprime que $\mathcal{A}$ est une algèbre de Lie pour l'opération $x, y \mapsto [x, y]$ (qui est visiblement linéaire en x et y). Cette algèbre sera appelée *algèbre de Lie des commutateurs* de l'algèbre associative $\mathcal{A}$ et notée $[\mathcal{A}]$.

Etant donné que tout homomorphisme d'algèbres associatives est manifestement un homomorphisme des algèbres des commutateurs correspondantes, la correspondance $\mathcal{A} \to [\mathcal{A}]$ est un foncteur de la catégorie ALG-ASS dans la catégorie ALG-LIE.

Il est souvent commode de désigner par $[x, y]$ le produit dans une algèbre de Lie quelconque (qui n'est l'algèbre de Lie des commutateurs d'aucune algèbre associative).

Dans ces notations, l'application ad a: $\mathfrak{g} \to \mathfrak{g}$, $\mathfrak{g}$ étant une algèbre de Lie quelconque, sera définie par la formule

$$(\operatorname{ad} a)\, x = [a, x], \quad a, x \in \mathfrak{g}.$$

Un exemple d'algèbre de Lie des commutateurs est l'algèbre de Lie des commutateurs $[\operatorname{End} \mathcal{V}]$ de l'algèbre associative $\operatorname{End} \mathcal{V}$ des endomorphismes (opérateurs linéaires) d'un espace vectoriel $\mathcal{V}$. Lorsque $\mathcal{V}$ est une algèbre (pas forcément associative), l'algèbre $[\operatorname{End} \mathcal{V}]$ contient le sous-espace vectoriel $\mathcal{D}(\mathcal{V})$ des *dérivations* de l'algèbre $\mathcal{V}$, c'est-à-dire des applications linéaires $D : \mathcal{V} \to \mathcal{V}$, telles que

$$D (xy) = Dx \cdot y + x \cdot Dy$$

pour tous $x, y \in \mathcal{V}$. Un calcul immédiat montre que pour tous D_1, $D_2 \in \mathcal{D} (\mathcal{V})$, le crochet de Lie $[D_1, D_2] = D_1D_2 - D_2D_1 \in \mathcal{D} (\mathcal{V})$, c'est-à-dire que l'espace vectoriel $\mathcal{D} (\mathcal{V})$ est une sous-algèbre de l'algèbre de Lie $[\operatorname{End} \mathcal{V}]$. Donc, pour *toute algèbre* $\mathcal{V}$, *l'espace vectoriel*

$\mathscr{D}(\mathscr{F})$ est une algèbre de Lie pour l'opération $D_1, D_2 \mapsto D_1D_2 - D_2D_1$.

Soit maintenant $\mathfrak{g}$ une algèbre de Lie arbitraire dont la loi de multiplication est désignée par $[x, y]$.

L'anticommutativité nous permet d'écrire l'identité de Jacobi (1) sous les formes suivantes:

$$[a, [x, y]] = [[a, x], y] + [x [a, y]], \qquad a, x, y \in \mathfrak{g},$$
$$[[a, b], x] = [a, [b, x]] - [b, [a, x]], \qquad a, b, x \in \mathfrak{g}.$$

La première de ces identités exprime que *pour tout élément $a \in \mathfrak{g}$ l'application*

$$(\mathrm{ad}\ a)\, x = [a, x], \quad x \in \mathfrak{g},$$

est une dérivation de l'algèbre de Lie $\mathfrak{g}$, la deuxième, que *l'application $a \mapsto$* ad *a de l'algèbre de Lie $\mathfrak{g}$ dans l'algèbre de Lie $\mathscr{D}(\mathfrak{g})$ est un homomorphisme.*

Les dérivations de la forme ad a s'appellent *dérivations intérieures* de l'algèbre de Lie $\mathfrak{g}$. On voit donc que *l'ensemble* ad $\mathfrak{g}$ *des dérivations intérieures d'une algèbre de Lie $\mathfrak{g}$ est une algèbre de Lie image de $\mathfrak{g}$ par un homomorphisme.*

En théorie des variétés différentiables, les algèbres de Lie (sur $\mathbb{R}$) apparaissent comme des algèbres de champs de vecteurs.

Soient M une variété différentiable arbitraire, $\mathfrak{a}(M)$, un espace vectoriel de champs de vecteurs sur M. On rappelle que tout champ $X \in \mathfrak{a}(M)$ peut être traité comme une dérivation sur M (un opérateur différentiel linéaire), c'est-à-dire comme une loi associant à tout ouvert $U \subset M$ la dérivation X_U de l'algèbre $\mathscr{F}(U)$ des fonctions différentiables sur U et telle que pour tout ouvert $V \subset U$ et toute fonction $f \in \mathscr{F}(U)$ l'on ait $X_V (f|_V) = (X_U f)|_V$. Donc, pour tous champs $X, Y \in \mathfrak{a}(M)$ et tout ouvert $U \subset M$ est définie la dérivation $[X_U, Y_U]$ de l'algèbre $\mathscr{F}(U)$. Les dérivations $[X_U, Y_U]$ forment un champ de vecteurs, puisqu'il est immédiat que pour tout ouvert $V \subset U$ et toute fonction $f \in \mathscr{F}(U)$, on a

$$[X_V, Y_V](f|_V) = ([X_U, Y_U] f)|_V.$$

Définition 3. Un champ de vecteurs sur M associant à tout ouvert $U \subset M$ la dérivation $[X_U, Y_U]$ de l'algèbre $\mathscr{F}(U)$ est dit *crochet de Lie* des champs X et Y et noté par le symbole $[X, Y]$. Donc, par définition

$$[X, Y]_U = [X_U, Y_U].$$

Il est évident que l'espace vectoriel $\mathfrak{a}(M)$ est une algèbre de Lie pour l'opération $X, Y \mapsto [X, Y]$. Cette algèbre s'appelle *algèbre de Lie des champs de vecteurs* sur la variété M. Cette algèbre est généralement de dimension infinie.

Un calcul immédiat nous montre que *dans toute carte* $(U, x^1, \ldots$ $\ldots, x^n)$ *les composantes* $[X, Y]^i$, $i = 1, \ldots, n$, *du champ* $[X, Y]$ *s'expriment en fonction des composantes* X^i, Y^i, $i = 1, \ldots, n$, *des champs* X *et* Y *à l'aide de la formule*

$$(2) \qquad [X, Y]^i = X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}, \qquad i, j = 1, \ldots, n.$$

En effet,

$$[X, Y]^i = [X, Y] x^i = X(Yx^i) - Y(Xx^i) =$$
$$= X(Y^i) - Y(X^i) = X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}. \quad \square$$

On montre avec la même aisance que si, Φ *étant une application différentiable de* M *sur* N, *des champs de vecteurs* X *et* Y *sur* M *sont* Φ*-liés respectivement à des champs* X' *et* Y' *sur* N, *le champ* $[X, Y]$ *l'est aussi à* $[X', Y']$. En effet, la Φ-liaison des champs X, Y et X', Y' exprime que pour toute fonction f (définie et différentiable sur un ouvert de la variété N), on a

$$X(f \circ \Phi) = X'f \circ \Phi \text{ et } Y(f \circ \Phi) = Y'f \circ \Phi.$$

Mais alors

$$[X, Y](f \circ \Phi) = X(Y(f \circ \Phi)) - Y(X(f \circ \Phi)) =$$
$$= X(Y'f \circ \Phi) - Y(X'f \circ \Phi) =$$
$$= X'(Y'f) \circ \Phi - Y'(X'f) \circ \Phi =$$
$$= [X', Y'] f \circ \Phi,$$

donc, les champs $[X, Y]$ et $[X', Y']$ sont Φ-liés aussi. $\square$

On constate, en particulier, que *pour tout difféomorphisme* Φ: $M \to N$, *on a*

$$\Phi^* [X, Y] = [\Phi^* X, \Phi^* Y],$$

où X, Y sont des champs de vecteurs arbitraires sur la variété N.

De là il s'ensuit immédiatement que *le crochet de Lie* $[X, Y]$ *de deux champs de vecteurs* X *et* Y *invariants à gauche sur un groupe de Lie* G *est aussi un champ de vecteurs invariant à gauche.* Cela signifie que l'espace vectoriel $\mathfrak{g} = \mathfrak{l}(G)$ des champs de vecteurs invariants à gauche est une sous-algèbre de l'algèbre de Lie $\mathfrak{a}(G)$ des champs de vecteurs, donc est lui-même une algèbre de Lie.

Définition 4. L'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$ sur $\mathbb{R}$ s'appelle *algèbre de Lie du groupe de Lie G.*

Signalons que nous avons construit le crochet de Lie $[X, Y]$ sur l'espace vectoriel $\mathfrak{l}(G)$ en nous plaçant dans le cadre de la première interprétation de cet espace. Il n'existe aucune construction immédiate dans le cadre de la deuxième interprétation ($\mathfrak{l}(G) = \mathsf{T}_e(G)$).

Nous verrons plus loin comment obtenir le crochet de Lie dans la troisième interprétation (c'est-à-dire lorsque les éléments de l'espace $\mathfrak{l}(G)$ sont traités comme des sous-groupes à un paramètre).

On sait que tout homomorphisme $\Phi: G \to H$ de groupes de Lie induit une application linéaire $\mathfrak{l}(\Phi): \mathfrak{l}(G) \to \mathfrak{l}(H)$ telle que pour tout champ de vecteurs $X \in \mathfrak{l}(G)$ le champ $\mathfrak{l}(\Phi)\, X \in \mathfrak{l}(H)$ est Φ-lié au champ X. Donc, cette application est un homomorphisme d'algèbres de Lie.

Ainsi, le foncteur de Lie $\mathfrak{l}$: GR-DIFF $\to$ LIN$_f(\mathbb{R})$ est bien un foncteur

$$\mathfrak{l}: \text{GR-DIFF} \to \text{ALG}_f\text{-LIE}$$

de la catégorie des groupes de Lie dans la catégorie ALG$_f$-LIE des algèbres de Lie sur $\mathbb{R}$ de dimension finie (plus exactement, le foncteur GR-DIFF $\to$ LIN$_f(\mathbb{R})$ est le composé du foncteur GR-DIFF $\to$ $\to$ ALG$_f$-LIE et du foncteur d'annihilation ALG$_f$-LIE $\to$ LIN$_f$ $(\mathbb{R})$).

Calculons le crochet de Lie pour les groupes de Lie de matrices. Etant donné que pour tout groupe de Lie de matrices G, l'algèbre $\mathfrak{l}(G)$ est visiblement une sous-algèbre de l'algèbre $\mathfrak{l}(\mathrm{GL}(n)) = \mathfrak{gl}(n)$, il nous suffit de calculer ce crochet seulement sur l'algèbre $\mathfrak{gl}(n)$.

Nous ferons les calculs directement pour un groupe arbitraire $G(\mathcal{A})$, où $\mathcal{A}$ est une algèbre associative de dimension finie.

On a établi dans la leçon 2 que l'espace vectoriel $\mathfrak{l}(G)$, où $G = G(\mathcal{A})$, s'identifie canoniquement à l'espace vectoriel $\mathcal{A}$, cette identification associant à un élément $a \in \mathcal{A}$ un champ de vecteurs invariant à gauche sur $G(\mathcal{A})$ de la forme $x \mapsto xa$, $x \in G(\mathcal{A})$ (ce champ est traité comme une application de $G(\mathcal{A})$ sur $\mathcal{A}$).

Soit $e_1, \ldots, e_n$ une base arbitraire de l'algèbre $\mathcal{A}$ et supposons comme dans la leçon 2 que

$$e_i e_j = c_{ij}^k e_k, \quad i, j, k = 1, \ldots, n,$$

où $c_{ij}^k \in \mathbb{R}$. Les coordonnées $x^1, \ldots, x^n$ des éléments de l'algèbre $\mathcal{A}$ dans la base $e_1, \ldots, e_n$ sont des coordonnées locales en tout point $x \in G(\mathcal{A})$ et de plus la base $\left(\frac{\partial}{\partial x^1}\right)_x, \ldots, \left(\frac{\partial}{\partial x^n}\right)_x$ de l'espace tangent $\mathsf{T}_x(G(\mathcal{A}))$ qui leur est associée est l'image de la base $e_1, \ldots, e_n$ par l'identification $\mathsf{T}_x(G(\mathcal{A})) = \mathcal{A}$. Donc, l'identification des champs de vecteurs sur $G(\mathcal{A})$ aux applications $G(\mathcal{A}) \to \mathcal{A}$ associe le champ de vecteurs $X = X^i \frac{\partial}{\partial x^i}$ à l'application $x \mapsto X^i e_i$.

Comme $xa = c_{jk}^i x^j a^k e_i$ pour $x = x^j e_j$, $a = a^k e_k$, il s'ensuit de là que le champ de vecteurs invariant à gauche $X: x \mapsto xa$ associé à $a \in \mathcal{A}$ admet dans la base $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$ les coordonnées

$$X^i = c_{jk}^i x^j a^k.$$

Donc, les coordonnées $[X, Y]^l$ du crochet de Lie de tous champs de vecteurs X et Y invariants à gauche sur $G(\mathcal{A})$ seront définies par la formule

$$[X, Y]^l = X^k \frac{\partial Y^l}{\partial x^k} - Y^k \frac{\partial X^l}{\partial x^k} =$$
$$= c^k_{ij} x^i a^j \cdot c^l_{km} b^m - c^k_{ij} x^i b^j \cdot c^l_{km} a^m =$$
$$= c^l_{km} (c^k_{ij} x^i a^j) b^m - c^l_{km} (c^k_{ij} x^i b^j) a^m,$$

où a et b sont les éléments de l'algèbre $\mathcal{A}$ correspondant aux champs X et Y. Or, cette formule exprime que les nombres $[X, Y]^l$ sont les coordonnées du point $xa \cdot b - xb \cdot a = x(ab - ba)$, donc au champ $[X, Y]$ est associé l'élément $ab - ba = [a, b]$. Ceci prouve qu'en vertu de l'identification $\mathfrak{l}(G(\mathcal{A})) = \mathcal{A}$, *l'algèbre de Lie des commutateurs* $[\mathcal{A}]$ *de l'algèbre associative* $\mathcal{A}$ *est l'algèbre de Lie du groupe de Lie* $G(\mathcal{A})$.

En particulier, *l'algèbre des commutateurs* $[\mathbb{R}(n)]$ *de l'algèbre des matrices* $\mathbb{R}(n)$ *est l'algèbre de Lie* $\mathfrak{gl}(n)$ *du groupe de Lie* $\mathrm{GL}(n)$.

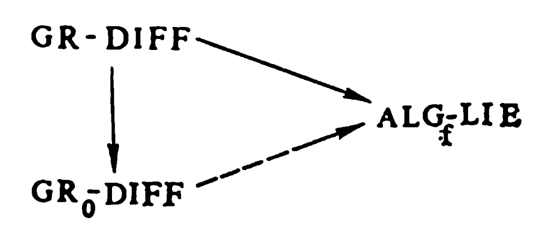
Pour tout groupe de matrices G, l'algèbre de Lie $\mathfrak{l}(G)$ est la sous-algèbre correspondante de l'algèbre de Lie $[\mathbb{R}(n)] = \mathfrak{gl}(n)$.

Notre prochain objectif est d'étudier en détail le foncteur de Lie $\mathfrak{l}$ et d'établir, en particulier, dans quelle mesure (et comment) ce foncteur est inversible, c'est-à-dire dans quelle mesure un groupe de Lie G peut être déterminé par son algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$.

Dans la leçon précédente, on a signalé que $\mathfrak{l}(G) = \mathfrak{l}(G_e)$, où G_e est la composante de l'unité du groupe G. En termes plus pédants, cette égalité affirme que l'injection $G_e \to G$ induit l'isomorphisme $\mathfrak{l}(G_e) \approx \mathfrak{l}(G)$. Mais puisqu'un homomorphisme de groupes de Lie induit un homomorphisme d'algèbres de Lie, l'isomorphisme $\mathfrak{l}(G_e) \approx \mathfrak{l}(G)$ sera aussi un isomorphisme d'algèbres de Lie. En d'autres termes, *l'égalité* $\mathfrak{l}(G_e) = \mathfrak{l}(G)$ *a lieu pour les algèbres de Lie.*

Donc, le problème de l'inversibilité du foncteur $\mathfrak{l}$: GR-DIFF $\to$ $\to$ ALG_f-LIE ne doit être posé que pour les groupes de Lie *connexes.*

On désignera par GR_0-DIFF la sous-catégorie complète de la catégorie GR-DIFF, engendrée par les groupes de Lie connexes. La restriction du foncteur de Lie à cette sous-catégorie sera aussi nommée *foncteur de Lie.* D'après ce qui précède on a le diagramme commutatif de foncteurs suivant:



dont les flèches obliques sont des foncteurs de Lie, et la flèche verticale, le foncteur de la composante de l'unité, qui à tout groupe de Lie associe la composante de son unité. Ce diagramme traduit formellement l'égalité $\mathfrak{l}(G_e) = \mathfrak{l}(G)$ dans le langage fonctoriel.

Le foncteur de Lie sera-t-il inversible sur la catégorie GR_0-DIFF ou, plus exactement, les groupes dont les algèbres de Lie sont isomorphes seront-ils isomorphes? Il s'avère que non.

Considérons par exemple le groupe additif $\mathbb{R}$ des réels et le groupe multiplicatif S^1 des nombres complexes de module 1. L'algèbre de ces deux groupes est de dimension un. Mais, en vertu de l'anticommutativité, la multiplication dans toute algèbre de Lie de dimension un (sur $\mathbb{R}$) est triviale (le produit de deux éléments quelconques est nul). Donc, les algèbres de Lie des groupes $\mathbb{R}$ et S^1 sont isomorphes, alors que ces groupes ne le sont pas (l'un étant compact, l'autre non).

Dans cet exemple, la cause de la coïncidence des algèbres de Lie est évidente : les groupes $\mathbb{R}$ et S^1 admettent une « même structure » locale (au voisinage de l'unité).

Ceci nous suggère d'introduire pour les groupes de Lie connexes la relation « d'isomorphisme local » en considérant que des groupes G et H sont *localement isomorphes* si un voisinage U de l'unité du groupe G peut être appliqué par un difféomorphisme sur un voisinage V de l'unité du groupe H, de telle sorte que le produit xy de deux éléments quelconques x, y de U se transforme, lorsque $xy \in U$, en le produit $\bar{x}\bar{y}$ des images $\bar{x}$, $\bar{y}$. Il est évident que *les algèbres de Lie de deux groupes de Lie localement isomorphes sont isomorphes* et l'on peut espérer (du moins sur le vu de l'exemple précédent) que, réciproquement, *les groupes de Lie dont les algèbres de Lie sont isomorphes sont localement isomorphes*. Il se trouve que c'est bien le cas. La démonstration de ce fait fondamental constituera l'un de nos principaux objectifs. Nous l'aborderons à la prochaine leçon et l'achèverons seulement dans la leçon 9.

Cependant la notion d'isomorphisme local de groupes de Lie « globaux » n'est pas très heureuse sur le plan méthodologique. L'expérience d'élaboration des théories mathématiques montre que le mélange du « global » et du « local » donne lieu à des formulations peu élégantes et complique inutilement les démonstrations. Il faut d'entrée de jeu tenter de séparer les problèmes globaux des locaux.

S'agissant des groupes de Lie, ces considérations générales plaident pour l'introduction d'une notion mathématique nouvelle, la notion de groupuscules de Lie, qui est une formalisation d'un voisinage de l'unité d'un groupe de Lie avec le produit défini sur ce voisinage.

Définition 5. On dit qu'une variété différentiable G est un *groupuscule de Lie* si

1) l'unité $e \in G$;

2) sont définis un voisinage $U \subset G \times G$ de l'élément (e, e) et un voisinage $U_0 \subset G$ de l'élément e;

3) sont définies une application différentiable

$$U \to G \tag{3}$$

appelée *multiplication* et une application différentiable

$$U_0 \to G \tag{4}$$

appelée *inversion*; l'image d'un point $(x, y) \in U$ par l'application (3) est désignée par xy et celle d'un point $x \in U_0$ par l'application (4), par x^{-1};

4) $e^{-1} = e$;

5) pour tout $(x, e) \in U$, on a $xe = x$ et pour tout $(e, x) \in U$, $ex = x$;

6) pour $(x, y) \in U$, $(y, z) \in U$, $(xy, z) \in U$ et $(x, yz) \in U$, on a

$$(xy)z = x(yz);$$

7) pour $(x, y) \in U$, $y \in U_0$ et $(xy, y^{-1}) \in U$, on a

$$(xy)y^{-1} = x,$$

et, de façon analogue, pour $(x, y) \in U$, $x \in U_0$ et $(x^{-1}, xy) \in U$, on a

$$x^{-1}(xy) = y.$$

Bref, G est un groupuscule de Lie si pour des éléments x, y assez voisins de l'unité e, sont définis le produit xy et l'élément réciproque x^{-1}, dépendant différentiablement, le premier de x et y, le second de x, les axiomes de groupe étant réalisés chaque fois que les objets figurant dans ces axiomes sont définis.

Les groupuscules de Lie sont également appelés *groupes de Lie locaux*.

Un exemple classique de groupuscule de Lie nous est donné par un voisinage quelconque de l'unité d'un groupe de Lie.

Même si elle semble évidente, la définition 5 est peu satisfaisante, car elle ne reflète pas tous les aspects de la notion intuitive que l'on se propose de formaliser. En effet, il semble logique d'admettre que deux voisinages distincts de l'unité d'un groupe de Lie conduisent à un même groupuscule de Lie, alors qu'aux termes de la définition 5 ces groupuscules seront différents.

Pour remédier à cette situation, on remarquera que tout ouvert H d'un groupuscule de Lie G, contenant l'unité e, est automatiquement un groupuscule de Lie. Un tel groupuscule s'appelle *sous-groupuscule de Lie*. Deux groupuscules de Lie sont dits *équivalents* si cer-

tains de leurs sous-groupuscules sont confondus. La classe des groupuscules de Lie équivalents s'appelle *germe* de groupuscules de Lie. (Comparer avec la définition des germes des fonctions différentiables en théorie des variétés différentiables.)

Les germes des groupuscules de Lie formalisent de façon adéquate la notion intuitive de groupe de Lie traité localement. Mais, l'usage des germes alourdit sensiblement l'exposé. En pratique, un style plus délié (que nous adopterons) prévaut et quand on parlera des groupuscules de Lie, on sous-entendra tacitement leurs germes; quant aux passages aux groupuscules équivalents, ils ne seront pas généralement mentionnés, ou, s'ils le seront, ce ne sera que dans les cas « cruciaux ».

On recommande vivement au lecteur de donner une forme plus pédante à la suite de l'exposé en faisant nettement la distinction entre les groupuscules de Lie et leurs germes.

Définition 6. On appelle *homomorphisme* d'un groupuscule de Lie G dans un groupuscule de Lie H une application différentiable Φ d'un voisinage V de l'unité de G sur le groupuscule H telle que

$$\Phi(xy) = \Phi x \cdot \Phi y,$$

pourvu que les éléments $\Phi(xy)$ et $\Phi x \cdot \Phi y$ soient définis. Si G_1 et H_1 sont des sous-groupuscules de G et de H et si $\Phi(V \cap G_1) \subset H_1$, alors Φ définit un homomorphisme du groupuscule G_1 dans le groupuscule H_1, appelé *partie de l'homomorphisme* Φ. Deux homomorphismes sont *équivalents* s'ils possèdent une partie commune. La classe des homomorphismes équivalents s'appelle *germe d'homomorphismes* (ou *homomorphisme de germes*).

Les groupuscules de Lie et leurs homomorphismes (plus exactement, les germes des groupuscules de Lie et leurs homomorphismes) forment de toute évidence une catégorie qui sera désignée par GR-LOC. Le passage à un voisinage quelconque de l'unité est manifestement défini par un foncteur

$$\text{GR-DIFF} \to \text{GR-LOC}$$

(de même que par un foncteur GR_0-DIFF $\to$ GR-LOC) de la catégorie GR-DIFF des groupes de Lie (resp. de la catégorie GR_0-DIFF des groupes de Lie connexes) dans la catégorie GR-LOC des groupuscules de Lie. Ce foncteur sera appelé *foncteur de localisation*. L'image d'un groupe de Lie G par un foncteur de localisation sera parfois notée G_{loc}. Les groupes de Lie sont localement isomorphes si leurs foncteurs de localisation sont isomorphes (comme objets de la catégorie GR-LOC).

Pour les groupuscules de Lie, on définit de manière évidente les champs de vecteurs invariants à gauche et leurs crochets de Lie. L'espace vectoriel $\mathfrak{l}(G)$ des champs de vecteurs invariants à gau-

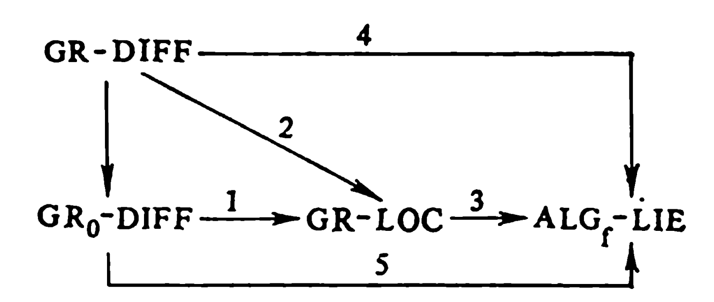
che sur un groupuscule de Lie G est une algèbre de Lie appelée *algèbre de Lie du groupuscule G.* De même que pour les groupes de Lie, les éléments de l'algèbre de Lie $\mathfrak{l}(G)$ s'identifient canoniquement aux vecteurs tangents $A \in \mathrm{T}_e(G)$ ainsi qu'aux sous-groupes à un paramètre (ou, pour être plus exact, aux sous-groupuscules à un paramètre, mais ce terme n'est pas usité) du groupuscule G.

Le foncteur

$$\mathfrak{l} : \text{GR-LOC} \to \text{ALG}_f\text{-Lie}$$

sera aussi nommé *foncteur de Lie.*

Les foncteurs introduits ci-dessus forment le diagramme commutatif suivant:



dont la flèche verticale gauche représente un foncteur associant à un groupe de Lie arbitraire sa composante de l'unité. les flèches *1* et *2*, des foncteurs de localisation, les flèches *3*, *4* et *5*, des foncteurs de Lie.

Nous voyons que le foncteur GR-DIFF → ALG_f-LIE qui nous intéresse en premier lieu est le composé de trois foncteurs: le foncteur GR-DIFF → GR_0-DIFF, le foncteur de localisation GR_0-DIFF → GR-LOC et le foncteur de Lie GR-LOC → ALG_f-LIE de groupuscules de Lie. Donc, l'étude du foncteur GR-DIFF → ALG_f-LIE se ramène à celle de ces trois foncteurs. L'étude de chacun d'eux implique des méthodes spéciales qui n'ont rien de commun entre elles. Nous avons ainsi atteint notre objectif qui était de discriminer les aspects locaux et les aspects globaux. La partie locale du problème est figurée par le foncteur GR-LOC → ALG_f-LIE, la partie globale, par les foncteurs GR-DIFF → GR_0-DIFF et GR_0-DIFF → GR-LOC.

Le foncteur GR-DIFF → GR_0-DIFF a été examiné dans la leçon 1. On y a prouvé que tout groupe de Lie G est l'extension de sa composante $H = G_e$ par un groupe discret. Réciproquement, toute extension G d'un groupe de Lie connexe H par un groupe discret est visiblement un groupe de Lie avec $G_e = H$. En première approximation ceci est décrit de manière satisfaisante par le foncteur GR-DIFF → GR_0-DIFF.

Remettons l'étude du foncteur GR_0-DIFF → GR-LOC à la leçon 10 et occupons-nous dans l'immédiat du foncteur $\mathfrak{l}$: GR-LOC → → ALG_f-LIE.

LEÇON 4

Exponentielle d'un opérateur différentiel linéaire.— Formule de calcul des valeurs de fonctions différentiables dans un voisinage normal de l'unité d'un groupe de Lie.— Formule de calcul des valeurs de fonctions différentiables sur le produit de deux éléments.— Série de Campbell-Hausdorff et polynômes de Dynkine.— Convergence de la série de Campbell-Hausdorff.— Détermination d'un groupuscule de Lie par son algèbre de Lie.— Opérations dans l'algèbre de Lie d'un groupe de Lie et sous-groupes à un paramètre.— Différentielles des automorphismes intérieurs.— Différentielle de l'application exponentielle.— Coordonnées canoniques.— Unicité de la structure de groupe de Lie.— Groupes sans petits sous-groupes et cinquième problème de Hilbert.

Soit M une variété différentiable. Nous supposerons dorénavant que M est une variété *analytique* (de classe C^{ω}). Puisque tout champ de vecteurs X sur M peut être traité comme un opérateur différentiel linéaire opérant sur des fonctions différentiables, on peut par analogie au cas matriciel introduire l'opérateur linéaire

$$(1) \qquad e^X = E + X + \frac{X^2}{2!} + \ldots + \frac{X^n}{n!} + \ldots = \sum_{n=0}^{\infty} \frac{X^n}{n!},$$

où E est l'opérateur identique, X^n, l'itération n-uple de l'opérateur X.

Certes, nous devons définir ce qu'on entend par somme de la série infinie (1). Pour cela il nous faut en principe munir l'espace $\mathcal{O}(M)$ d'une topologie (par exemple, au moyen d'une norme). Mais, pour simplifier, nous aborderons cette question de biais et entendrons par convergence de la série (4) la convergence « faible ». Plus exactement, nous d é f i n i r o n s l'action $e^X f$ de l'opérateur e^X sur une

fonction $f \in \Theta(M)$ par la formule

$$(2) \qquad e^X f = f + Xf + \frac{X^2 f}{2!} + \ldots + \frac{X^n f}{n!} + \ldots = \sum_{n=0}^{\infty} \frac{X^n f}{n!},$$

où $X^n f = X(X^{n-1} f)$, et nous admettrons que cet opérateur n'agit que sur les fonctions dont la série (2) (qui est déjà fonctionnelle) admet un domaine de convergence non vide (domaine qui sera pris pour domaine de définition de la fonction $e^X f$).

L'opérateur e^X ne sera plus un opérateur différentiel. Plus exactement, nous allons montrer qu'en général il est induit par un difféomorphisme $\Phi: M \to M$, c'est-à-dire est de la forme $f \mapsto f \circ \Phi$. A cet effet, il suffit seulement que les courbes intégrales $t \mapsto \varphi_a(t)$ du champ X soient « assez longues », c'est-à-dire qu'elles soient définies pour $|t| \leqslant 1$ et que le domaine de définition $W(f)$ de la fonction f soit « assez grand », c'est-à-dire qu'il existe des points $a \in W(f)$ tels que $\varphi_a(t) \in W(f)$ pour $|t| \leqslant 1$. La fonction $e^X f$ sera alors définie pour tous les points tels que a et s'exprimera par

$$(3) \qquad (e^X f)(a) = f(\varphi_a(1)).$$

En effet, introduisons la fonction

$$F(t) = f(\varphi_a(t)).$$

Par hypothèse, cette fonction est définie et analytique pour $|t| \leqslant 1$. Elle se développe donc en la série de Taylor

$$F(t) = \sum_{n=0}^{\infty} \frac{F^{(n)}(0)}{n!} t^n,$$

qui converge pour $|t| \leqslant 1$. Par ailleurs, puisque la courbe $t \mapsto \varphi_a(t)$ est une courbe intégrale du champ X, il vient

$$F'(t) = \frac{df(\varphi_a(t))}{dt} = \frac{d\varphi_a(t)}{dt} f = X_{\varphi_a(t)} f = (Xf)(\varphi_a(t)).$$

Ceci exprime que la fonction $F(t)$ de la fonction Xf n'est autre que la fonction $F'(t)$, d'où par une récurrence immédiate il s'ensuit que la fonction $F(t)$ de la fonction $X^n f$ n'est autre que la fonction $F^{(n)}(t)$, c'est-à-dire que

$$F^{(n)}(t) = (X^n f)(\varphi_a(t)).$$

Donc

$$F^{(n)}(0) = (X^n f)(\varphi_a(0)) = (X^n f)(a),$$

et, par suite,

$$F(t) = \sum_{n=0}^{\infty} \frac{(X^n f)(a)}{n!} t^n = \sum_{n=0}^{\infty} \frac{(tX)^n f}{n!}(a) = (e^{tX} f)(a).$$

Pour achever la démonstration, il reste à poser $t = 1$. □

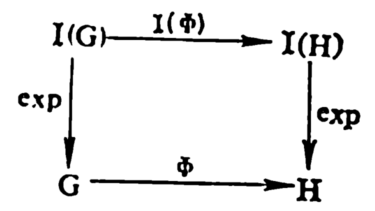
Nous appliquerons la formule générale (3) aux champs de vecteurs invariants à gauche X sur un groupe (ou un groupuscule) de Lie analytique et aux fonctions f définies et analytiques dans un voisinage de l'unité e du groupe G. Pour point a nous prendrons l'unité e. Désignons le point $\varphi_e(1)$ par exp X et écrivons pour ce cas la formule (3) sous la forme suivante:

(4) $$f(\exp X) = (e^X f)(e).$$

C'est cette forme que nous utiliserons dans la suite.

La courbe intégrale $t \mapsto \varphi_e(t)$ est un sous-groupe à un paramètre $t \mapsto \beta(t)$ pour le champ de vecteurs invariant à gauche X. Donc, $\exp X = \beta(1)$ et la formule (4) est valable pour toutes fonctions f dont l'ensemble de définition contient la portion $t \to \beta(t)$, $|t| \leqslant 1$, de ce sous-groupe.

Par définition, exp est une application de l'espace vectoriel $\mathfrak{g} = \mathfrak{l}(G)$ dans le groupe G telle que $\exp 0 = e$. Elle est visiblement canonique, c'est-à-dire que pour tout homomorphisme $\Phi: G \to H$ de groupes de Lie, on a le diagramme commutatif



L'application exp: $\mathfrak{l}(G) \to G$ est différentiable en vertu du théorème de dépendance différentiable des solutions d'une équation différentielle par rapport aux conditions initiales.

Assertion A. *L'application*

$$\exp: \mathfrak{l}(G) \to G$$

est un difféomorphisme au point 0.

Ce fait sera établi ultérieurement.

Définition 1. On dit qu'un voisinage $\mathring{U}$ du vecteur nul $0 \in \mathfrak{g}$ est *normal* si

a) il est é t o i l é, c'est-à-dire qu'il contient tout vecteur X avec les vecteurs tX, où $|t| \leqslant 1$;

b) l'application exp est un difféomorphisme du voisinage $\mathring{U}$ sur un voisinage U de l'unité e du groupe G.

Le voisinage U sera aussi appelé *voisinage normal.*

Aux termes de l'assertion A, il existe des voisinages normaux (aussi bien du vecteur $0 \in \mathfrak{g}$ que de l'unité $e \in G$) qui sont aussi petits que l'on veut (c'est-à-dire qui peuvent être compris dans tout voisinage donné à priori).

Par construction, $\exp X = \beta(1)$, où β est un sous-groupe à un paramètre, est une courbe intégrale du champ X. Or il est clair que la courbe $t \mapsto \beta(at)$ qui est visiblement un sous-groupe à un paramètre, est une courbe intégrale du champ aX, où $a \in \mathbb{R}$. Donc, $\exp(aX) = \beta(a)$ ou, si l'on désigne a par t,

$$\exp(tX) = \beta(t).$$

Cette formule exprime que $\beta: t \to \exp(tX)$, c'est-à-dire que *les sous-groupes à un paramètre* β *du groupe* G *sont les images des droites* $t \mapsto tX$ *par l'application* exp.

En vertu de la condition a) de la définition 1, il s'ensuit de là que la condition imposée au domaine de définition de la fonction f, condition qui est nécessaire (et suffisante) pour que la formule (4) soit valable, sera visiblement remplie si ce domaine est un voisinage normal U du point e. Sous ces conditions, puisque tout point de U est de la forme $\exp X$, où $X \in \mathring{U}$, *la formule* (4) *définit la fonction* f *sur le voisinage* U *tout entier.*

Sans perdre ceci de vue, appliquons la formule (4) au calcul de la valeur $f(ab)$ de la fonction f sur le produit ab des éléments $a = \exp X$ et $b = \exp Y$ (sous réserve bien sûr que $ab \in U$).

Pour un élément $a \in U$, la formule $f_a(b) = f(ab)$ définit (sur un voisinage du point e supposé normal compris dans U) une fonction f_a différentiable. De plus, si $b = \exp Y$, la formule (4) nous donne $f_a(b) = (e^Y f_a)(e)$.

D'autre part, la fonction f_a n'étant autre que la composée $f \circ L_a$ de la translation à gauche $L_a: b \mapsto ab$ et de la fonction f, il s'ensuit, en vertu de l'invariance à gauche du champ Y, que

$$Yf_a = Yf \circ L_a = (Yf)_a.$$

Mais alors $Y^n f_a = (Y^n f)_a$ pour tout $n \geqslant 0$, et, par suite,

$$e^Y f_a = (e^Y f)_a.$$

Donc, si $a = \exp X$, en appliquant de nouveau la formule (4), on trouve

$$f_a(b) = (e^Y f_a)(e) = (e^Y f)(a) = (e^X e^Y f)(e).$$

Comme $f_a(b) = f(\exp X \cdot \exp Y)$, ceci prouve que

$$f(\exp X \cdot \exp Y) = (e^X e^Y f)(e) \tag{5}$$

pour tous X et Y du voisinage normal correspondant du zéro de l'algèbre $\mathfrak{g}$.

Par définition (nous faisons provisoirement abstraction de la convergence)

$$e^X e^Y = \sum_{p=0}^{\infty} \frac{1}{p!} X^p \left(\sum_{q=0}^{\infty} \frac{1}{q!} Y^q \right) = \sum_{p,\, q=0}^{\infty} \frac{1}{p!\, q!} X^p Y^q.$$

En portant cette série dans la série logarithmique

$$\ln Z = (Z-E) - \frac{(Z-E)^2}{2} + \ldots = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (Z-E)^k,$$

on obtient (puisque les opérateurs X et Y ne commutent généralement pas) la série formelle

$$\ln (e^X e^Y) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left(\sum_{\substack{p,\, q=0 \\ p+q>0}}^{\infty} \frac{X^p Y^q}{p!\, q!} \right)^k =$$

$$= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum \frac{X^{p_1} Y^{q_1} \ldots X^{p_k} Y^{q_k}}{p_1!\, q_1! \ldots p_k!\, q_k!},$$

où la sommation dans la somme intérieure est étendue à tous les ensembles $(p_1, \ldots, p_k, q_1, \ldots, q_k)$ des entiers positifs tels que

$$p_1 + q_1 > 0, \ \ldots, \ p_k + q_k > 0. \tag{6}$$

En posant

$$z_n(x,\, y) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \sum \frac{x^{p_1} y^{q_1} \ldots x^{p_k} y^{q_k}}{p_1!\, q_1! \ldots p_k!\, q_k!}, \tag{7}$$

où dans la somme intérieure les indices $p_1, \ldots, p_k, q_1, \ldots, q_k$ vérifient à la fois la condition (6) et la condition suivante

$$p_1 + \ldots + p_k + q_1 + \ldots + q_k = n, \tag{8}$$

on peut mettre la série $\ln (e^X e^Y)$ sous la forme suivante:

$$\ln (e^X e^Y) = \sum_{n=1}^{\infty} z_n(X,\, Y). \tag{9}$$

Cette série formelle s'appelle *série de Campbell-Hausdorff*.

Du point de vue algébrique, $z_n(x, y)$ n'est autre qu'un polynôme de variables x et y généralement non permutables. Etudions ces polynômes plus en détail.

Soient x et y des symboles, K, un corps arbitraire. On forme des polynômes en x et en y sur K en effectuant des opérations d'addition et de multiplication autant de fois qu'on le veut sur x, y et les élé-

ments de K. On construit de façon analogue des *polynômes non commutatifs* en x et en y sur K; la seule différence est que leur multiplication n'est pas supposée être commutative (mais comme toujours $kf = fk$, $\forall f$ et $\forall k \in K$). Ces polynômes forment une algèbre unitaire que nous désignerons par $K\langle x, y\rangle$. (Cette algèbre sera définie de façon plus formelle dans la leçon suivante.)

Les *polynômes de Lie* en x et en y obtenus par les opérations d'addition et de multiplication sur des éléments de K ainsi que par le crochet de Lie $a, b \mapsto [a, b] = ab - ba$ se définissent de façon naturelle dans l'algèbre des commutateurs $[K\langle x, y\rangle]$ de l'algèbre $K\langle x, y\rangle$. Il est clair qu'ils forment une sous-algèbre de l'algèbre de Lie $[K\langle x, y\rangle]$, qui contient les éléments x et y et qui est ellemême contenue dans une telle sous-algèbre (ou, en termes algébriques, est engendrée par les éléments x et y). Nous désignerons cette sous-algèbre par le symbole $\mathfrak{l}\langle x, y\rangle$.

Chaque polynôme de Lie $u \in \mathfrak{l}\langle x, y\rangle$ est aussi un polynôme de $K\langle x, y\rangle$ (il suffit de développer le crochet de Lie). En tant que tel le polynôme u sera désigné par ιu. Formellement, l'application

$$\iota : \mathfrak{l}\langle x, y\rangle \to K\langle x, y\rangle, \quad u \mapsto \iota u,$$

est définie comme une application linéaire telle que $\iota[a, b] = \iota a \cdot \iota b - \iota b \cdot \iota a$ pour tous éléments $a, b \in \mathfrak{l}\langle x, y\rangle$. Elle est injective par définition.

Si le corps K est de caractéristique 0, la formule (7) définit un élément $z_n(x, y)$ dans $K\langle x, y\rangle$.

Assertion B. *Il existe un polynôme de Lie* د$_n(x, y)$ *tel que*

$$\iota\, د_n(x, y) = z_n(x, y).$$

Le symbole د représente la lettre arabe « dāl ».

Nous prouverons l'assertion B ultérieurement.

Exemples.

Supposons que $n = 1$. Il est évident que

$$z_1(x, y) = x + y,$$

et, par suite,

$$د_1(x, y) = x + y$$

(de sorte que le crochet de Lie ne participe pas dans la construction du polynôme د$_1(x, y)$).

Supposons $n = 2$. Pour $k = 1$, la somme intérieure de la formule (7) contient trois termes: $\frac{1}{2}x^2$, xy et $\frac{1}{2}y^2$ et pour $k = 2$, quatre termes: x^2, xy, yx et y^2. Donc

$$z_2(x, y) = \frac{1}{2}xy - \frac{1}{2}yx,$$

et, par suite,

$$\mathfrak{z}_2(x, y) = \frac{1}{2}[x, y].$$

On vérifie de façon analogue que

$$z_3(x, y) = \frac{1}{12}x^2y + \frac{1}{12}yx^2 + \frac{1}{12}xy^2 + \frac{1}{12}y^2x - \frac{1}{6}xyx - \frac{1}{6}yxy.$$

Il est difficile d'appréhender directement la forme du polynôme $\mathfrak{z}_3(x, y)$ (et on se demande même s'il existe). Cependant par quelques calculs on s'assure que

$$\mathfrak{z}_3(x, y) = \frac{1}{12}[x, [x, y]] + \frac{1}{12}[y, [y, x]].$$

Les calculs croissent très rapidement avec n. Ceci n'empêche qu'il est possible d'établir pour les polynômes $\mathfrak{z}_n(x, y)$ une formule explicite identique à la formule (7) du polynôme $z_n(x, y)$. Cette formule ayant été acquise pour la première fois par E. Dynkine, les polynômes $\mathfrak{z}_n(x, n)$ seront nommés *polynômes de Dynkine.*

A noter que *le polynôme $\mathfrak{z}_n(x, n)$ est homogène de degré n par rapport à x et à y*, c'est-à-dire que

$$\mathfrak{z}_n(tx, ty) = t^n \mathfrak{z}_n(x, y)$$

pour tout t.

Puisque l'opération $X, Y \mapsto [X, Y]$ s'exprime aussi par la formule $[X, Y] = XY - YX$ pour les champs de vecteurs (les opérateurs différentiels linéaires), de l'assertion B il s'ensuit que *pour tous opérateurs $X, Y \in \mathfrak{l}(G)$, l'opérateur $z_n(X, Y)$ appartient à l'algèbre* $\mathfrak{l}(G)$. Donc, l'algèbre $\mathfrak{l}(G)$ contient aussi l'opérateur $\ln(e^Xe^Y)$ à condition bien sûr que celui-ci ait un sens, c'est-à-dire que la série (9) converge. Voyons donc la convergence de cette série ou, plus exactement, de la série

$$(10) \qquad \mathfrak{z}_1(X, Y) + \mathfrak{z}_2(X, Y) + \ldots + \mathfrak{z}_n(X, Y) + \ldots,$$

de polynômes de Lie $\mathfrak{z}_n(X, Y)$ en X et en Y.

Tous les termes de la série (10) appartiennent à l'espace vectoriel $\mathfrak{g} = \mathfrak{l}(G)$, donc on aurait pu étudier la convergence de cette série pour une norme (par exemple euclidienne) de $\mathfrak{g}$.

Cependant les majorations indispensables impliquent une information détaillée sur la structure des polynômes $\mathfrak{z}_n(x, y)$, une information qui ne nous sera accessible qu'à la leçon suivante. Force est donc d'opter pour une autre voie et de revenir à la série (9), dont la structure des termes nous est connue, quitte à perdre l'avantage de la finitude de la dimension.

Aux termes de notre conception générale de la convergence des séries d'opérateurs, on comprendra la convergence de la série (9)

dans un sens « faible », mais un peu plus fort que plus haut. Plus exactement, on admettra que l'opérateur ln $(e^X e^Y)$ défini par la série (9) agit sur une fonction $f \in \mathcal{O}(G)$ si la série fonctionnelle

$$(11) \quad z_1(X, Y) f + z_2(X, Y) f + \ldots + z_n(X, Y) f + \ldots$$

converge dans l'ensemble de définition de f. La somme g de cette série sera le résultat de l'application de l'opérateur ln $(e^X e^Y)$ à la fonction f.

Donc, en vertu de cette définition, on a $W(g) = W(f)$ (alors qu'avant $W(g) \subset W(f)$).

Nous allons montrer, dans l'hypothèse que $\mathfrak{g}$ est munie d'une norme $\| \quad \|$, qu'*il existe un nombre* $\delta_0 > 0$ *tel que pour* $\| X \| < \delta_0$ *et* $\| Y \| < \delta_0$ *l'opérateur* ln $(e^X e^Y)$ *s'applique à toute fonction* $f \in \mathcal{O}(G)$. Puisque tout élément de G qui appartient à l'ensemble de définition de la fonction f possède un voisinage U d'adhérence $\bar{U}$ compacte appartenant aussi à cet ensemble, il suffit pour cela de prouver que *pour tout voisinage* U *d'adhérence* $\bar{U}$ *compacte, l'opérateur* ln$(e^X e^Y)$ *est défini sur l'espace vectoriel* $\mathcal{F}(\bar{U})$ *des fonctions différentiables sur* $\bar{U}$.

On remarquera à cet effet que puisque l'ensemble $\bar{U}$ est compact, pour toute fonction $f \in \mathcal{F}(\bar{U})$ est défini le nombre

$$\| f \| = \max \left(|f|, \left| \frac{\partial f}{\partial x^1} \right|, \ldots, \left| \frac{\partial f}{\partial x^n} \right| \right),$$

où $x^1, \ldots, x^n$ sont les coordonnées locales dans un voisinage U (ou, plus exactement, dans un voisinage plus grand contenant l'adhérence $\bar{U}$ de U). Le nombre $\| f \|$ dépend évidemment du choix des coordonnées $x^1, \ldots, x^n$, mais cette circonstance n'a aucune incidence sur la suite. On vérifie immédiatement que la fonctionnelle $f \mapsto \| f \|$ est une norme sur l'espace vectoriel $\mathcal{F}(\bar{U})$. Ceci étant, puisque $Xf = X^i \frac{\partial f}{\partial x^i}$ et $| X^i | \leqslant \| X^i \|$, $\forall i$, alors tout champ $X \in \mathfrak{g}$ tel que $\| X \| < \delta$ est justiciable de la majoration

$$\| Xf \| \leqslant \delta \| f \|.$$

De là on déduit par une récurrence immédiate que si $\| X \| < \delta$ et $\| Y \| < \delta$, pour tous indices $p_1, \ldots, p_k, q_1, \ldots, q_k$, on a la majoration

$$\| X^{p_1} Y^{q_1} \ldots X^{p_k} Y^{q_k} f \| \leqslant \delta^n \| f \|,$$

où $n = p_1 + \ldots + p_k + q_1 + \ldots + q_k$. Cela signifie que la série (11) est majorée (après une division par $\| f \|$) par la série ln $(e^X e^Y)$ dans laquelle X et Y sont remplacés par δ et les coefficients, par leurs valeurs absolues, c'est-à-dire est majorée par la

série déduite de

$$\sum_{n=1}^{\infty} \frac{(z-1)^n}{n} \tag{12}$$

par la substitution à z de la série

$$e^{2\delta} = \sum_{n=0}^{\infty} \frac{(2\delta)^n}{n!}.$$

Comme la série (12) converge pour $|z-1|<1$, cette série majorante convergera aussi pour $|e^{2\delta}-1|<1$, c'est-à-dire pour $\delta < \frac{\ln 2}{2}$. Donc, pour $\delta < \frac{\ln 2}{2}$, la série (11) converge uniformément dans $\bar{U}$ vers une fonction différentiable, c'est-à-dire que l'opérateur $\ln(e^X e^Y)$ s'applique à la fonction f. □

Comme la série (10) ne diffère de la série (9) que par les notations (par définition $z_n(X, Y) = \mathfrak{z}_n(X, Y)$ pour tous champs $X, Y \in \mathfrak{g}$), ceci prouve que pour $\|X\| < \delta_0$, $\|Y\| < \delta_0$, où $\delta_0 = \frac{\ln 2}{2}$, la série (10) converge au même sens « faible ». Or, on sait que pour les séries d'opérateurs dont les termes appartiennent à un espace vectoriel de dimension finie, la convergence « faible » coïncide avec la convergence « forte ». Donc, *la série* (10) *converge fortement pour* $\|X\| < \delta_0$ *et* $\|Y\| < \delta_0$.

Remarque 1. On peut, sans faire référence au théorème de coïncidence des convergences « faible » et « forte », démontrer la convergence de la série (10) en considérant un système arbitraire de coordonnées locales $x^1, \ldots, x^n$ au point e du groupe G. Pour $f = x^i$, $i = 1, \ldots, n$, la série (11) se transforme en une série des composantes $z_n(X, Y)^i = \mathfrak{z}_n(X, Y)^i$ des champs de vecteurs $\mathfrak{z}_n(X, Y)$. Donc, cette série converge. Par suite, il en est de même de la série de leurs valeurs $\mathfrak{z}_n(X, Y)_e^i$ au point e, c'est-à-dire de la série des composantes des vecteurs $\mathfrak{z}_n(X, Y)_e$ $\left(\text{dans la base } \left(\frac{\partial}{\partial x^1}\right)_e, \ldots, \left(\frac{\partial}{\partial x^n}\right)_e\right)$. Ceci exprime la convergence de la série de vecteurs

$$\mathfrak{z}_1(X, Y)_e + \mathfrak{z}_2(X, Y)_e + \ldots + \mathfrak{z}_n(X, Y)_e + \ldots \tag{13}$$

Soient $\mathfrak{z}(X, Y)_e$ la somme de cette série et $\mathfrak{z}(X, Y)$ un champ de vecteurs invariant à gauche prenant en e la valeur $\mathfrak{z}(X, Y)_e$. Pour tout élément $a \in G$, la série

$$\mathfrak{z}_1(X, Y)_a + \mathfrak{z}_2(X, Y)_a + \ldots + \mathfrak{z}_n(X, Y)_a + \ldots$$

des valeurs prises par les termes de la série (10) en e se déduit de la série (13) par action de l'opérateur linéaire dL_a, donc elle converge

vers le vecteur $(dL_a)(\mathfrak{z}(X, Y)_e) = \mathfrak{z}(X, Y)_a$. Ce qui exprime que la série (10) converge vers $\mathfrak{z}(X, Y)$.

En Analyse, on démontre que si une fonction se développe en une série entière convergente, cette série est unique. De là il s'ensuit, en particulier, que si dans la série de e^z on porte la série de $\ln z$, on obtient, après réduction des termes semblables, la série de z. En d'autres termes, l'égalité $e^{\ln z} = z$ est valable aussi bien pour les fonctions que pour les séries formelles. Donc, elle reste valable dans le cas aussi où l'on y remplace z, par exemple, par le série de $e^X e^Y$. Cela signifie que si dans la série de e^z on remplace z par la série (10), on obtient la série de $e^X e^Y$. Comme toutes ces séries convergent, ceci prouve que

$$e^{\mathfrak{z}(X,Y)} = e^X e^Y. \tag{14}$$

Cette formule a lieu pour tous éléments X, Y de l'algèbre $\mathfrak{l}(G)$ tels que $\| X \| < \delta_0$ et $\| Y \| < \delta_0$, où δ_0 est un nombre arbitraire assez petit (d'après l'étude faite plus haut tout $\delta_0 < \frac{\ln 2}{2}$ convient).

Revenons maintenant à la formule (5). Cette formule a lieu pour tous X et Y d'un voisinage normal $\mathring{U}$ du zéro de l'algèbre $\mathfrak{g} = \mathfrak{l}(G)$, tels que $\exp X \cdot \exp Y \in U = \exp \mathring{U}$. Il existe en particulier un δ_1 tel que pour tout δ strictement positif $< \delta_1$, la formule (5) est valable pour $\| X \| < \delta$ et $\| Y \| < \delta$. Donc, en admettant de plus que $\delta < \delta_0$, on peut mettre la formule (5) sous la forme suivante:

$$f(\exp X \cdot \exp Y) = (e^Z f)(e), \text{ où } Z = \mathfrak{z}(X, Y).$$

Remarquons maintenant qu'un élément Z de l'algèbre $\mathfrak{l}(G)$ dépend continûment des éléments X et Y, de sorte que $Z \to 0$ lorsque $X \to 0$ et $Y \to 0$. Donc, pour δ assez petit, le champ Z est justiciable de la formule (4) qui dit que

$$(e^Z f)(e) = f(\exp Z).$$

Donc, si $\| X \| < \delta$ et $\| Y \| < \delta$, où $\delta > 0$ est assez petit, alors

$$f(\exp X \cdot \exp Y) = f(\exp Z)$$

pour toute fonction différentiable f définie aux points $\exp X \cdot \exp Y$ et $\exp Z$. Ceci est en particulier vrai si f est l'une des coordonnées $x^1, \ldots, x^n$ en e. Donc, les coordonnées $x^1, \ldots, x^n$ prennent toutes les mêmes valeurs aux points $\exp X \cdot \exp Y$ et $\exp Z$, ce qui n'est possible que si

$$\exp X \cdot \exp Y = \exp Z.$$

En résumé, on obtient le théorème suivant qui est la finalité de nos raisonnements précédents.

Théorème 1. *L'unité e d'un groupe de Lie (ou d'un groupuscule) analytique possède un voisinage U jouissant des propriétés suivantes:*

a) *il existe un* $\delta > 0$ *tel que chaque point du voisinage U se représente d'une seule manière par* $\exp X$, *où* $X \in \mathfrak{l}(G)$ *et* $\| X \| < \delta$;

b) *pour tous points* $\exp X$ *et* $\exp Y$ *de U, il existe dans l'algèbre* $\mathfrak{l}(G)$ *un élément Z tel que*

$$\exp X \cdot \exp Y = \exp Z; \tag{15}$$

c) *cet élément Z est la somme* $\mathfrak{Z}(X, Y)$ *de la série* (10) *dont les termes sont les polynômes de Dynkine* $\mathfrak{Z}_n(X, Y)$ *en X et en Y.* □

Ce théorème dit qu'un groupe (groupuscule) de Lie G possède une partie sur laquelle la multiplication est déterminée de façon unique (conformément à la formule (15)) par l'algèbre de Lie $\mathfrak{l}(G)$. Donc, *deux groupuscules de Lie (analytiques) dont les algèbres de Lie sont isomorphes, sont isomorphes* (plus exactement, ont des germes isomorphes). Dans ce sens on peut dire que *le foncteur de Lie*

$$L\colon \text{GR-LOC} \to \text{ALG}_f\text{-LIE}$$

est inversible à un isomorphisme près.

On établira une proposition plus exacte à la leçon 6.

Le théorème 1 nous permet de résoudre aisément le problème, qui a été ajourné, de l'interprétation des opérations de l'algèbre $\mathfrak{g} = \mathfrak{l}(G)$ en termes de sous-groupes à un paramètre.

Si dans l'espace $\mathfrak{g}$ on choisit arbitrairement une base $e_1, \ldots, e_n$, alors pour tout voisinage normal U de l'unité du groupe G, le composé h du difféomorphisme $\exp^{-1}\colon U \to \mathring{U}$ et de la restriction de l'isomorphisme $\mathfrak{g} \to \mathbb{R}^n$ à $\mathring{U}$ sera un difféomorphisme de U sur un ouvert de l'espace $\mathbb{R}^n$, c'est-à-dire que le couple (U, h) sera une carte sur le groupe de Lie G. Les coordonnées locales correspondantes s'appellent *coordonnées normales.* Donc, si $a = \exp X$ et $X = x^1 e_1 + \ldots + x^n e_n$, les nombres $x^1, \ldots, x^n$ sont les coordonnées normales du point $a \in U$.

On désignera par β_X le sous-groupe à un paramètre $t \mapsto \exp(tX)$ correspondant à l'élément $X \in \mathfrak{g}$ (resp. la courbe intégrale du champ X qui passe par e pour $t = 0$ si X est traité comme un champ de vecteurs invariant à gauche; resp. le sous-groupe à un paramètre admettant X pour vecteur tangent en $t = 0$ si X est traité comme un vecteur de l'espace $\mathsf{T}_e(G)$).

Proposition 1. *Pour tout* $X \in \mathfrak{g}$ *et tout* $k \in \mathbb{R}$, *l'élément* $kX \in \mathfrak{g}$ *(traité comme un élément de l'espace* $\mathsf{T}_e(G)$*) est un vecteur tangent en* $t = 0$ *à la courbe*

$$t \mapsto \beta_X(kt). \tag{16}$$

Pour tous $X, Y \in \mathfrak{g}$, l'élément $X + Y \in \mathfrak{g}$ est (dans la même interprétation) un vecteur tangent, pour $t = 0$, à la courbe

$$t \mapsto \beta_X(t)\, \beta_Y(t), \tag{17}$$

et l'élément $[X, Y] \in \mathfrak{g}$, un vecteur tangent en $t = 0$ à la courbe

$$t \mapsto \beta_X(\sqrt{t})\, \beta_Y(\sqrt{t})\, \beta_X(\sqrt{t})^{-1} \beta_Y(\sqrt{t})^{-1}. \tag{18}$$

D é m o n s t r a t i o n. La première assertion est évidente, puisque la courbe (16), c'est-à-dire la courbe $t \mapsto \exp(ktX)$ n'est autre que le sous-groupe à un paramètre β_{kX}.

Pour prouver la deuxième assertion, on remarquera que puisque $\mathfrak{H}_n(tX, tY) = t^n \mathfrak{H}_n(X, Y)$ et $\mathfrak{H}_1(X, Y) = X + Y$, il vient

$$\mathfrak{H}(tX, tY) = t(X + Y) + O(t^2),$$

où $O(t^2)$ désigne les termes dont la puissance de t est $\geqslant 2$. Donc, la courbe (17) est de la forme

$$t \mapsto \exp(t(X + Y) + O(t^2)),$$

et, par suite, elle est définie par les fonctions

$$x^i(t) = t(X^i + Y^i) + O(t^2)$$

en coordonnées normales (dans une base arbitraire de l'algèbre de Lie $\mathfrak{g}$). Donc, le vecteur qui lui est tangent en $t = 0$ a pour composantes

$$\left.\frac{dx^i(t)}{dt}\right|_{t=0} = X^i + Y^i$$

et, par suite, est confondu avec le vecteur $X + Y$.

De façon analogue, puisque

$$\begin{aligned}(\exp X)\cdot(\exp Y)\cdot(\exp X)^{-1}\cdot(\exp Y)^{-1} &= \\ &= \exp X \cdot \exp Y \cdot \exp(-X) \cdot \exp(-Y) = \\ &= \exp(\mathfrak{H}(X, Y)) \exp(\mathfrak{H}(-X, -Y)) = \\ &= \exp \mathfrak{H}(\mathfrak{H}(X, Y), \mathfrak{H}(-X, -Y))\end{aligned}$$

et

$$\begin{aligned}\mathfrak{H}(\mathfrak{H}(X, Y), \mathfrak{H}(-X, -Y)) &= \mathfrak{H}(X, Y) + \mathfrak{H}(-X, -Y) + \\ &+ \frac{1}{2}[\mathfrak{H}(X, Y), \mathfrak{H}(-X, -Y)] + \ldots = \\ &= \left\{(X + Y) + \frac{1}{2}[X, Y] + \ldots\right\} + \\ &+ \left\{(-X, -Y) + \frac{1}{2}[-X, -Y] + \ldots\right\} + \\ &+ \frac{1}{2}[X + Y + \ldots, -X - Y + \ldots] + \ldots = \frac{1}{2}[X, Y] + \\ &+ \frac{1}{2}[X, Y] + \frac{1}{2}[X, -Y] + \frac{1}{2}[Y, -X] + \ldots = [X, Y] + \ldots,\end{aligned}$$

la courbe (18) prend la forme

$$t \mapsto \exp([\sqrt{t}\, X, \sqrt{t}\, Y] + O(t^{3/2})) = \exp(t[X, Y] + O(t^{3/2})),$$

et, par suite, elle est définie par les fonctions

$$(19) \qquad x^i(t) = t[X, Y]^i + O(t^{3/2}), \qquad i = 1, \ldots, n,$$

en coordonnées normales. Donc

$$\left.\frac{dx^i(t)}{dx}\right|_{t=0} = [X, Y]^i,$$

et, par suite, le vecteur tangent à la courbe (18) pour $t = 0$ est le vecteur $[X, Y]$. □

Remarque 2. Remarquons que la courbe (16) est un sous-groupe à un paramètre contrairement aux courbes (17) et (18). Bien plus, la courbe (18) n'est définie que pour $t \geqslant 0$, de sorte qu'on ne peut parler de son vecteur tangent pour $t = 0$. C'est pourquoi nous devons donner une définition *ad hoc* du vecteur tangent à la courbe (18) pour $t = 0$. Nous admettrons que ce vecteur est la limite des vecteurs tangents à la courbe (18) pour $t > 0$ lorsque $t \to 0$. De la formule (19) il s'ensuit que cette limite existe et est égale à $[X, Y]$.

Remarque 3. Il est utile de savoir que dans les formules (17) et (18), les sous-groupes à un paramètre β_X et β_Y peuvent être remplacés par des courbes arbitraires $\alpha_X: t \mapsto \alpha_X(t)$ et $\alpha_Y: t \mapsto \alpha_Y(t)$ passant pour $t = 0$ par le point e et admettant en ce point les vecteurs tangents X et Y. En effet, pour $|t|$ assez petit, dans $\mathfrak{g}$ seront définis des vecteurs $\mathring{\alpha}_X(t) = \exp^{-1}\alpha_X(t)$ et $\mathring{\alpha}_Y(t) = \exp^{-1}\alpha_Y(t)$, tels que

$$\mathring{\alpha}_X(t) = tX + O(t^2), \qquad \mathring{\alpha}_Y(t) = tY + O(t^2)$$

et, par suite,

$$\mathfrak{Z}(\mathring{\alpha}_X(t), \mathring{\alpha}_Y(t)) = t(X + Y) + O(t^2).$$

Donc, le vecteur tangent, par exemple à la courbe

$$t \mapsto \alpha_X(t)\,\alpha_Y(t) = \exp \mathfrak{Z}(\mathring{\alpha}_X(t), \mathring{\alpha}_Y(t)),$$

sera le vecteur $X + Y$ pour $t = 0$.

Le théorème 1 permet aussi de calculer la différentielle d'un automorphisme intérieur arbitraire.

Pour tout élément a du groupe de Lie G, la différentielle $(d\Phi_a)_e = \iota(\Phi_a)$ de l'automorphisme intérieur correspondant $\Phi_a: x \mapsto axa^{-1}$, $x \in G$, est notée $\mathrm{Ad}(a)$. C'est une application linéaire inversible

$$\mathrm{Ad}(a): \mathfrak{g} \to \mathfrak{g}$$

de l'espace vectoriel $\mathfrak{g} = \mathfrak{l}(G)$ sur lui-même. Comme $\Phi_a = L_a \circ R_{a^{-1}}$, il vient (dans l'interprétation $\mathfrak{g} = \mathsf{T}_e(G)$)

$$\operatorname{Ad}(a) = (dL_a^{-1})_{a^{-1}} \circ (dR_{a^{-1}})_e.$$

Il est clair que l'application

$$\operatorname{Ad}: a \mapsto \operatorname{Ad}(a)$$

est un homomorphisme du groupe de Lie G dans le groupe de Lie $\operatorname{Aut} \mathfrak{g}$ des opérateurs linéaires non dégénérés de l'espace vectoriel $\mathfrak{g}$. Cet homomorphisme s'appelle *représentation adjointe* du groupe de Lie G.

La différentielle $(d \operatorname{Ad})_e = \mathfrak{l}(\operatorname{Ad})$ de l'homomorphisme Ad au point e est une application linéaire de l'espace vectoriel $\mathfrak{g} = \mathfrak{l}(G)$ dans l'espace vectoriel $\operatorname{End} \mathfrak{g} = \mathfrak{l}(\operatorname{Aut} \mathfrak{g})$ des opérateurs linéaires de $\mathfrak{g}$. D'autre part, on sait de la leçon 3 que pour toute algèbre de Lie $\mathfrak{g}$, il existe une application linéaire canonique $\operatorname{ad}: \mathfrak{g} \to \operatorname{End} \mathfrak{g}$ définie par la formule

$$\operatorname{ad} X: Y \mapsto [X, Y], \qquad X, Y \in \mathfrak{g}.$$

Proposition 2. *On a l'égalité*

$$\mathfrak{l}(\operatorname{Ad}) = \operatorname{ad}.$$

Démonstration (comparer avec celle de la proposition 1). Comme

$$\begin{aligned} ᑐ(ᑐ(X, Y), -X) &= ᑐ\left(X + Y + \tfrac{1}{2}[X, Y] + \ldots, -X\right) = \\ &= Y + \tfrac{1}{2}[X, Y] + \tfrac{1}{2}[Y, -X] + \ldots = \\ &= Y + [X, Y] + \ldots, \end{aligned}$$

où les points de suspension désignent les termes de puissance $\geqslant 3$ en X et en Y, il vient

$$\begin{aligned} (\exp X)(\exp Y)(\exp X)^{-1} &= \exp ᑐ(ᑐ(X, Y), -X) = \\ &= \exp(Y + [X, Y] + \ldots), \end{aligned}$$

donc

$$\begin{aligned} \Phi_{\beta_X(s)}(\beta_Y(t)) &= (\exp(sX))(\exp(tY))(\exp(sX))^{-1} = \\ &= \exp(tY + st[X, Y] + \ldots), \end{aligned}$$

où les points de suspension représentent les termes de puissance $\geqslant 3$ en s et en t. Donc, les composantes normales du vecteur

$$(d\Phi_{\beta_X(s)})_e Y = (d\Phi_{\beta_X(s)})\left(\frac{d\beta_Y(t)}{dt}\bigg|_{t=0}\right) = \left(\frac{d}{dt}\Phi_{\beta_X(s)}(\beta_Y(t))\right)\bigg|_{t=0}$$

sont égales à

$$\frac{d}{dt}(tY^i + st[X, Y]^i + \ldots)|_{t=0} = Y^i + s[X, Y]^i + \ldots,$$

où les points de suspension du second membre représentent les termes de puissance $\geqslant 2$ en s, et, par suite,

$$(d\Phi_{\beta_X(s)})_e Y = Y + s[X, Y] + \dots$$

Comme

$$(d\,\mathrm{Ad})_e X = (d\,\mathrm{Ad})_e \left(\frac{d\beta_X(s)}{ds}\Big|_{s=0}\right) = \frac{d}{ds}\mathrm{Ad}(\beta_X(s))|_{s=0} =$$
$$= \lim_{s\to 0}\frac{\mathrm{Ad}(\beta_X(s)) - \mathrm{Ad}(e)}{s} = \lim_{s\to 0}\frac{(d\Phi_{\beta_X(s)})_e - E}{s},$$

il s'ensuit de là que

$$((d\,\mathrm{Ad})_e X)\,Y = \lim_{s\to 0}\frac{(d\Phi_{\beta_X(s)})_e Y - Y}{s} =$$
$$= \lim_{s\to 0}([X, Y] + \dots) = [X, Y] =$$
$$= (\mathrm{ad}\, X)\,Y. \quad \square$$

Corollaire. *Pour tout élément* $X \in \mathfrak{g}$, *on a*

$$\mathrm{Ad}(\exp X) = e^{\mathrm{ad}\, X}.$$

D é m o n s t r a t i o n. Les formules

$$t \mapsto \mathrm{Ad}(t \exp X) \quad \text{et} \quad t \mapsto e^{t\,\mathrm{ad}\, X}$$

définissent des sous-groupes à un paramètre du groupe de Lie Aut $\mathfrak{g}$ possédant, pour $t = 0$, le même vecteur tangent

$$(d\,\mathrm{Ad})_e X = \mathrm{ad}\, X,$$

donc ces sous-groupes sont confondus pour tous les t. $\square$

On obtient un autre exemple d'application du théorème 1 en considérant une courbe différentiable arbitraire $t \mapsto X(t)$ dans l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$. Pour tout $s \in \mathbb{R}$, l'application $t \mapsto \exp(sX(t))$ est une courbe dans G passant pour $t = 0$ par le point $a(s) = \exp(sX)$, où $X = X(0)$. Soit

$$A(s) = \frac{d}{dt}\exp(sX(t))\Big|_{t=0}$$

le vecteur tangent à cette courbe au point $a(s)$. En transportant ce vecteur au point e au moyen de la différentielle $(dR_{a(s)^{-1}})_{a(s)}$, on obtient dans $\mathfrak{g} = \mathsf{T}_e(G)$ le vecteur

$$B(s) = (dR_{a(s)^{-1}})_{a(s)}A(s).$$

L'application $s \mapsto B(s)$ est une courbe différentiable dans $\mathfrak{g}$, donc pour tout s est défini son vecteur tangent $B'(s)$. Il se trouve que

$$B'(s) = \mathrm{Ad}(a(s))\,Y, \tag{20}$$

où $Y = X'(0)$. En effet, par définition de l'action de la différentielle d'une application différentiable sur les vecteurs tangents aux courbes, on a

$$B(s) = \frac{d}{dt}\left(R_{a(s)^{-1}}\left(\exp\left(sX(t)\right)\right)\right)\Big|_{t=0} =$$
$$= \frac{d}{dt}\left(\exp\left(sX(t)\right)\right)\exp\left(-sX\right)\Big|_{t=0},$$

et, par suite, la remarque 3 nous donne

$$B(s+\Delta s) - B(s) = \frac{d}{dt}\left(\left(\exp\left(sX(t)\right)\exp\left(-sX\right)\right)^{-1} \times\right.$$
$$\times \left(\exp\left((s+\Delta s)X(t)\right)\exp\left(-(s+\Delta s)X\right)\right)|_{t=0} =$$
$$= \frac{d}{dt}\left(\exp\left(sX\right)\exp\left(-sX(t)\right) \times\right.$$
$$\times \exp\left((s+\Delta s)X(t)\right)\exp\left(-(s+\Delta s)X\right)\right)|_{t=0} =$$
$$= \frac{d}{dt}\left(a(s)\exp\left(\Delta s\, X(t)\right)a(s+\Delta s)^{-1}\right) =$$
$$= \left(dL_{a(s)} \circ dR_{a(s+\Delta s)^{-1}}\right)\frac{d}{dt}\left(\exp\left(\Delta s\, X(t)\right)\right)\Big|_{t=0}.$$

Donc

$$B'(s) = \lim_{\Delta s \to 0}\frac{B(s+\Delta s) - B(s)}{\Delta s} =$$
$$= \left(dL_{a(s)} \circ dR_{a(s)^{-1}}\right)\lim_{\Delta s \to 0}\frac{\frac{d}{dt}\left(\exp\left(\Delta s\, X(t)\right)\right)|_{t=0}}{\Delta s} =$$
$$= \mathrm{A}d\left(a(s)\right)\lim_{\Delta s \to 0}\frac{\frac{d}{dt}\left(\exp\left(\Delta s\, X(t)\right)\right)|_{t=0}}{\Delta s}$$

et pour prouver l'égalité (20), il suffit simplement de montrer que

$$\lim_{\Delta s \to 0}\frac{\frac{d}{dt}\left(\exp\left(\Delta s\, X(t)\right)\right)\Big|_{t=0}}{\Delta s} = Y.$$

Or, ceci est évident, puisque dans une base $e_1, \ldots, e_n$ de l'espace vectoriel $\mathfrak{g}$, le point $\exp(\Delta s\, X(t))$ a pour coordonnées normales $\Delta s\, X^i(t)$, où $X^i(t)$ sont les composantes du vecteur $X(t)$ dans la base $e_1, \ldots, e_n$, donc, le vecteur

$$\lim_{\Delta s \to 0}\frac{\frac{d}{dt}\left(\exp\left(\Delta s\, X(t)\right)|_{t=t_0}\right.}{\Delta s}$$

possède dans cette base les composantes $\frac{dX^i(0)}{dt}$, c'est-à-dire les mêmes composantes que le vecteur $X'(0) = Y$. □

L'opérateur linéaire Ad $(a(s))$ de la formule (20) peut être mis sous la forme suivante:

$$\mathrm{Ad}\,(a(s)) = \mathrm{Ad}\,(\exp(sX)) = e^{s\,\mathrm{ad}\,X} =$$
$$= E + s\,\mathrm{ad}\,X + \ldots + s^n \frac{(\mathrm{ad}\,X)^n}{n!} + \ldots$$

En intégrant cette identité par rapport à s entre 0 et 1, on obtient l'identité

$$\int_0^1 \mathrm{Ad}\,(a(s))\,ds = \frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X},$$

où $\dfrac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}$ désigne la somme de la série d'opérateurs

$$E + \frac{\mathrm{ad}\,X}{2!} + \ldots + \frac{(\mathrm{ad}\,X)^n}{(n+1)!} + \ldots,$$

déduite de la série entière de la fonction $\dfrac{e^z - 1}{z}$ par substitution de l'opérateur ad X à z. Pour l'opérateur $B(1)$, il s'ensuit de là, grâce à la formule (20), que

$$B(1) = \int_0^1 B'(s)\,ds = \frac{e^{\mathrm{ad}} - E}{\mathrm{ad}\,X}\,Y.$$

Puisque, par définition

$$B(1) = (dR_{a(1)^{-1}})_{a(1)} \frac{d}{dt} \exp X(t)\Big|_{t=0} =$$
$$= (dR_{\exp(-X)})_{\exp X} \frac{d}{dt} \exp X(t)\Big|_{t=0} =$$
$$= \frac{d}{dt}(\exp X(t) \exp(-X))\Big|_{t=0},$$

ceci prouve que

$$\frac{d}{dt}(\exp X(t) \exp(-X))\Big|_{t=0} = \frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}\,Y.$$

En passant aux coordonnées normales, on déduit immédiatement de là que le vecteur $Z(t) \in \mathfrak{g}$ qui vérifie la relation $\exp X(t) \exp(-X) = \exp Z(t)$ est justiciable de l'égalité

$$Z'(0) = \frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}\,Y,$$

d'où il résulte que

$$Z(t) = t\,\frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}\,Y + O(t^2).$$

En revenant à $\exp Z(t)$ et en posant $X(t) = X + tY$, on constate qu'on a démontré la

Proposition 3. *Pour tous éléments $X, Y \in \mathfrak{g}$, on a l'égalité*

$$\exp(X+tY)\exp(-X)=\exp\left(t\,\frac{e^{\operatorname{ad} X}-E}{\operatorname{ad} X}\,Y+O(t^2)\right). \ \square$$

Pour tout élément $X \in \mathfrak{g}$, la différentielle $(d\exp)_X$ au point X de l'application différentiable $\exp: \mathfrak{g} \to G$ est, en vertu de l'identification $\mathsf{T}_X(\mathfrak{g}) = \mathfrak{g}$, une application linéaire $\mathfrak{g} \to \mathsf{T}_a(G)$, où $a = \exp X$. Donc, la composée de cette différentielle et de l'application $(dR_a)_e^{-1}: \mathsf{T}_a(G) = \mathsf{T}_e(G) = \mathfrak{g}$ est une application de $\mathfrak{g}$ dans $\mathfrak{g}$.

Corollaire 1. *On a la formule*

$$(dR_a)_e^{-1}\circ(d\exp)_X=\frac{e^{\operatorname{ad} X}-E}{\operatorname{ad} X},\quad a=\exp X.$$

D é m o n s t r a t i o n. Soit $Y \in \mathfrak{g}$. Comme $\left.\frac{d(X+tY)}{dt}\right|_{t=0} = Y$, il vient

$$((dR_a^{-1})_e\circ(d\exp)_X)\,Y=\frac{d}{dt}(\exp(X+tY)\exp(-X))\Big|_{t=0}=$$
$$=\frac{d}{dt}\left(t\,\frac{e^{\operatorname{ad} X}-E}{\operatorname{ad} X}\,Y+O(t^2)\right)\Big|_{t=0}=$$
$$=\frac{e^{\operatorname{ad} X}-E}{\operatorname{ad} X}\,Y. \ \square$$

Corollaire 2. *L'application* $\exp: \mathfrak{g} \to G$ *est un difféomorphisme en un point* $X \in \mathfrak{g}$ *si et seulement si aucune racine caractéristique de l'opérateur* $\operatorname{ad} X$ *n'est de la forme* $2m\pi i$.

D é m o n s t r a t i o n. L'application exp est un difféomorphisme en X si et seulement si sa différentielle $(d\exp)_X$ en X est un isomorphisme, et l'opérateur $\operatorname{ad} X$ admet des racines caractéristiques de la forme $2m\pi i$ si et seulement si l'opérateur $e^{\operatorname{ad} X} - E$, donc l'opérateur $\frac{e^{\operatorname{ad} X}-E}{\operatorname{ad} X}$, est dégénéré. $\square$

Signalons que ces résultats ont été acquis sous l'hypothèse que les assertions A et B sont vraies. Aussi notre objectif immédiat sera-t-il de les prouver.

Nous commencerons par prouver l'assertion A sous une forme plus générale.

Supposons que l'espace vectoriel $\mathfrak{g} = \mathfrak{l}(G)$ a été décomposé en une somme directe

$$\mathfrak{g} = \mathcal{A} \oplus \mathcal{B} \tag{21}$$

de sous-espaces $\mathscr{A}$ et $\mathscr{B}$. Définissons l'application

(22) $$\Phi: \mathfrak{g} \to G,$$

en posant

$$\Phi(X) = \exp A \cdot \exp B, \quad \forall X \in \mathfrak{g},$$

où $A \in \mathscr{A}$ et $B \in \mathscr{B}$ sont les composantes de $X \in \mathfrak{g}$ dans la décomposition (21). Il est évident que cette application est différentiable et associe au zéro de $\mathfrak{l}(G)$ l'unité e de G. Calculons la différentielle

$$(d\Phi)_0: \mathsf{T}_0(\mathfrak{g}) \to \mathsf{T}_e(G)$$

de cette application en 0.

Soit

$$l: \mathfrak{g} \to \mathsf{T}_0(\mathfrak{g})$$

l'isomorphisme canonique associant à $X \in \mathfrak{g}$ le vecteur tangent en 0 à la courbe $t \mapsto tX$. L'application Φ associe à cette courbe la courbe

(23) $$t \mapsto \exp tA \cdot \exp tB = \exp \mathfrak{Z}(tA, tB),$$

et, par suite, sa différentielle $(d\Phi)_0$ associe au vecteur $l(X)$ le vecteur tangent en e à la courbe (23). Ceci signifie que pour toute fonction $f \in \mathcal{O}_e(G)$, on a la formule

$$[((d\Phi)_0 \circ l)(X)] f = \left.\frac{df(\exp \mathfrak{Z}(tA, tB))}{dt}\right|_{t=0},$$

et, par suite (cf. formule (4)), la formule

$$[((d\Phi)_0 \circ l) X)] f = \left.\frac{d\left(e^{\mathfrak{Z}(tA, tB)} f\right)(e)}{dt}\right|_{t=0}.$$

Or

$$e^{\mathfrak{Z}(tA, tB)} f = (E + \mathfrak{Z}(tA, tB) + O(t^2)) f = \\ = f + t(A+B)f + O(t^2),$$

donc

$$\left.\frac{d\left(e^{\mathfrak{Z}(tA, tB)} f\right)(e)}{dt}\right|_{t=0} = ((A+B)f)(e) = (Xf)(e) = X_e f.$$

Par conséquent,

$$((d\Phi)_0 \circ l)(X) = X_e,$$

c'est-à-dire que

$$(d\Phi)_0 \circ l = i,$$

où i est un isomorphisme $X \to X_e$ de l'espace vectoriel $\mathfrak{g} = \mathfrak{l}(G)$ sur l'espace vectoriel $\mathsf{T}_e(G)$.

Comme i et l sont des isomorphismes, il s'ensuit de là qu'il en est de même de l'application $(d\Phi)_0$. En vertu du théorème classique

des applications étales (localement difféomorphes), ceci prouve la proposition suivante:

Proposition 4. *L'application* (22) *est un difféomorphisme au point* $0 \in \mathfrak{g}$. ☐

Pour $\mathfrak{g} = \mathcal{A}$ (et $\mathcal{B} = 0$), on obtient l'assertion A qui de ce fait est entièrement prouvée.

En vertu de la proposition 4, le point $0 \in \mathfrak{g}$ possède un voisinage étoilé $\mathring{U}$ aussi petit que l'on veut sur lequel l'application (22) est un difféomorphisme de $\mathring{U}$ sur un voisinage U de l'unité e de G.

Définition 2. Les voisinages $\mathring{U}$ et U jouissant de cette propriété sont dits *voisinages canoniques* (des points $0 \in \mathfrak{g}$ et $e \in G$ respectivement) associés à la décomposition directe (21).

En choisissant des bases arbitraires dans les sous-espaces $\mathcal{A}$ et $\mathcal{B}$, on obtient une base de l'espace $\mathfrak{g}$. Le composé h du difféomorphisme $\Phi^{-1}: U \to \mathring{U}$ et de la restriction de l'isomorphisme $\mathfrak{g} \to \mathbb{R}^n$ à $\mathring{U}$ est un difféomorphisme du voisinage U sur un ouvert de l'espace $\mathbb{R}^n$, c'est-à-dire que le couple (U, h) sera une carte sur G.

De telles cartes sont dites *cartes canoniques* associées à la décomposition (21), et les coordonnées locales $x^1, \ldots, x^n$, *coordonnées canoniques.*

Il est clair que toutes ces définitions (de même que la proposition 4) se généralisent immédiatement au cas où

$$\mathfrak{g} = \mathcal{A}_1 \oplus \ldots \oplus \mathcal{A}_m. \tag{24}$$

Si $\mathcal{A} = \mathfrak{g}$ et $\mathcal{B} = 0$ dans la décomposition (21), c'est-à-dire si $m = 1$ dans la décomposition (24), les voisinages canoniques coïncident avec les voisinages normaux au sens de la définition 1, et les coordonnées canoniques, avec les coordonnées normales.

L'autre cas extrême se présente pour $m = n$ lorsque les espaces $\mathcal{A}_1, \ldots, \mathcal{A}_m$ sont tous de dimension un (c'est-à-dire lorsque la décomposition (24) est définie par une base dans $\mathfrak{g}$). Les coordonnées canoniques correspondantes s'appellent *coordonnées canoniques de deuxième espèce* (et les coordonnées normales, *coordonnées canoniques de première espèce*).

On remarquera que les coordonnées canoniques de première et de deuxième espèce sont définies par une base arbitraire de l'espace vectoriel $\mathfrak{g}$.

La proposition importante suivante se démontre sans peine à l'aide des coordonnées canoniques:

Proposition 5. *Tout homomorphisme continu* Φ *de groupes de Lie* H *et* G (*traités comme des groupes topologiques*) *est une application différentiable* (*un homomorphisme de groupes de Lie*).

D é m o n s t r a t i o n. Soient $a \in H$ et $g \in \mathcal{O}_{\Phi(a)}(G)$. Il nous faut montrer que la fonction $g \circ \Phi$ définie au voisinage du point a est différentiable en a, c'est-à-dire appartient à $\mathcal{O}_a(H)$. Mais, puisque Φ est un homomorphisme, on a

$$\Phi = L_{\Phi a}^{-1} \circ \Phi \circ L_a.$$

Donc, l'application L_a et la fonction $f = g \circ L_{\Phi a}^{-1}$ étant différentiables, il nous suffit de prouver que pour toute fonction $f \in \mathcal{O}_e(G)$, la fonction $f \circ \Phi$ est différentiable dans un voisinage de $e \in H$, c'est-à-dire que l'application Φ est différentiable en e.

Traitons d'abord le cas où H est le groupe additif $\mathbb{R}$ des réels. Soient U un voisinage normal (canonique de première espèce) de l'unité $e \in G$, $x^1, \ldots, x^n$, les coordonnées normales correspondantes. Soit, par ailleurs, $\varepsilon > 0$ un nombre tel que $\Phi(t) \in U$ lorsque $|t| < \varepsilon$. Si $0 < t < \varepsilon$ et $0 < |r| < s$, où r et s sont des entiers, alors

$$\Phi\left(\frac{r}{s} t\right) = \Phi\left(\frac{t}{s}\right)^r \quad \text{et} \quad \Phi(t) = \Phi\left(\frac{t}{s}\right)^s.$$

D'autre part, les coordonnées $x^1, \ldots, x^n$ étant normales, pour tout $i = 1, \ldots, n$, tout $a \in U$ et tout s tel que $a^s \in U$, on a

$$x^i(a^s) = sx^i(a)$$

(il suffit de remarquer que si $a = \exp A$, alors $a^s = \exp(sA)$). Donc

$$(x^i \circ \Phi)\left(\frac{r}{s} t\right) = r \cdot (x^i \circ \Phi)\left(\frac{t}{s}\right)$$

et

$$(x^i \circ \Phi)(t) = s \cdot (x^i \circ \Phi)\left(\frac{t}{s}\right).$$

Par suite,

$$(x^i \circ \Phi)\left(\frac{r}{s} t\right) = \frac{r}{s} \cdot (x^i \circ \Phi)(t).$$

L'application Φ étant continue, on a l'égalité analogue

$$(x^i \circ \Phi)(\alpha t) = \alpha \cdot (x^i \circ \Phi)(t)$$

pour tout réel α, $|\alpha| < 1$. Ceci exprime que les fonctions $x^i \circ \Phi$ sont linéaires. Donc, l'application Φ est définie en coordonnées locales par des fonctions linéaires, donc analytiques. Par conséquent, elle est différentiable.

Supposons maintenant que le groupe H est arbitraire. Considérons une base quelconque $Y_1, \ldots, Y_n$ de son algèbre de Lie $\mathfrak{h}$. Alors pour tout $i = 1, \ldots, n$, l'application $t \mapsto \Phi(\exp tY_i)$ sera un homomorphisme $\mathbb{R} \to G$ continu, et, par suite, différentiable, en vertu de ce qui a été déjà prouvé, c'est-à-dire un sous-groupe à un

paramètre du groupe G. Donc, dans l'algèbre de Lie $\mathfrak{g}$ du groupe G, il existe des éléments $X_1, \ldots, X_n$ tels que

$$\Phi(\exp tY_i) = \exp tX_i$$

pour tout $i = 1, \ldots, n$. Comme Φ est un homomorphisme, il s'ensuit de là que pour tous $t^1, \ldots, t^n \in \mathbb{R}$ on a

$$\Phi(\exp t^1Y_1 \ldots \exp t^nY_n) = \exp t^1X_1 \ldots \exp t^nX_n.$$

Il est clair que l'élément $\exp t^1X_1 \ldots \exp t^nX_n$ du groupe G dépend différentiablement de $t^1, \ldots, t^n$, c'est-à-dire que ses coordonnées locales $x^1, \ldots, x^n$ (dans une carte quelconque) sont des fonctions différentiables $x^1(t), \ldots, x^n(t)$ de $t = (t^1, \ldots, t^n)$. Or, par définition, les nombres $t^1, \ldots, t^n$ ne sont autres (dans un voisinage de l'unité du groupe H) que les coordonnées canoniques de deuxième espèce définies par la base considérée de l'algèbre $\mathfrak{h}$. Donc, les fonctions $x^1(t), \ldots, x^n(t)$ sont des fonctions définissant l'application Φ dans les coordonnées locales $t^1, \ldots, t^n$ et $x^1, \ldots, x^n$. Ces fonctions étant différentiables, il en est de même de l'application Φ. □

Corollaire. *Si deux groupes de Lie sont isomorphes en tant que groupes topologiques, ils le sont en tant que groupes de Lie.* □

De là il s'ensuit, en particulier, que *si l'on peut munir un groupe topologique d'une structure différentiable compatible avec sa topologie de telle sorte qu'il devienne un groupe de Lie, on ne peut le faire que d'une seule manière.*

Cela signifie que le foncteur d'annihilation

$$\text{GR-DIFF} \to \text{GR-TOP}$$

associe à des groupes de Lie distincts des groupes topologiques distincts. On peut donc considérer que ce foncteur réalise une *injection* de la catégorie GR-DIFF dans la catégorie GR-TOP. En d'autres termes, la catégorie GR-DIFF des groupes de Lie peut être traitée comme une sous-catégorie de la catégorie GR-TOP des groupes topologiques. La proposition 5 exprime alors que cette sous-catégorie est complète.

Il est naturel de se poser maintenant la question de savoir s'il n'est pas possible de caractériser la sous-catégorie GR-DIFF à l'intérieur de la catégorie GR-TOP par des conditions topologiques générales sans faire intervenir la différentiabilité, c'est-à-dire s'il n'est pas possible, à l'intérieur de la catégorie GR-TOP, de caractériser les groupes topologiques G admettant une structure de groupes de Lie.

Une condition nécessaire est évidente: un groupe topologique G admettant une structure de groupe de Lie doit nécessairement être

séparé et localement compact. Pour formuler une condition nécessaire plus fine. nous aurons besoin de la définition suivante:

Définition 3. On dit qu'un groupe topologique G est *sans petits sous-groupes* si son unité e possède un voisinage ne contenant aucun sous-groupe $H \neq \{e\}$.

Il se trouve que cette propriété est nécessaire pour que le groupe G puisse être muni d'une structure de groupe de Lie.

Proposition 6. *Tout groupe de Lie G (ou, plus exactement, tout groupe topologique G_{top} déduit du groupe G en ignorant la différentiabilité) est un groupe sans petits sous-groupes.*

D é m o n s t r a t i o n. Munissons l'espace $\mathsf{T}_e(G) = \mathfrak{g}$ d'une métrique euclidienne arbitraire. Alors, pour $\delta > 0$ assez petit, la boule de rayon δ centrée en 0 sera un voisinage normal de 0, et, par suite, son image par l'application exp sera un voisinage normal de l'unité du groupe G. Soit U un voisinage normal construit comme ci-dessus pour $\delta/2$. Il est clair que pour tout vecteur $A \in \mathfrak{g}$ non nul de module $<\delta/2$, il existe un entier m tel que le module du vecteur mA soit compris entre $\delta/2$ et δ. Cela signifie que pour tout élément $a = \exp A$ de U différent de l'unité, il existe un m tel que $a^m = \exp mA$ n'appartienne pas à U. Donc, le voisinage U ne peut contenir aucun sous-groupe $H \neq \{e\}$. □

Il est remarquable qu'avec la condition de compacité locale, l'absence de petits sous-groupes est une condition suffisante pour qu'un groupe séparé G puisse être muni d'une structure de groupe de Lie (structure qui est unique comme déjà prouvé).

Théorème (Glisson-Yamabe). *Un groupe séparé topologique est un groupe de Lie si et seulement s'il est localement compact et ne possède pas de petits sous-groupes.*

La démonstration du théorème de Glisson-Yamabe est trop compliquée pour être développée ici.

Une autre condition nécessaire pour qu'un groupe topologique soit un groupe de Lie est qu'il soit *localement euclidien*, c'est-à-dire qu'il soit une variété topologique. La question de savoir si cette condition est suffisante constitue le *cinquième problème de Hilbert* (dans sa formulation moderne). Une étude minutieuse de la structure des groupes topologiques possédant de petits sous-groupes a conduit (Montgomery, Tsippine, Iwasawa. Glisson, Yamabe) à la conclusion qu'*aucun groupe localement euclidien ne peut contenir de petits sous-groupes*. Combiné au théorème de Glisson-Yamabe ceci nous donne immédiatement une réponse positive au problème de Hilbert: *tout groupe localement euclidien est un groupe de Lie.*

Voir les détails dans [5] et [7].

LEÇON 5

Algèbres associatives libres.— Algèbres de Lie libres.—Lemme fondamental.— Algèbre enveloppante universelle.— Injection d'une algèbre de Lie dans son algèbre enveloppante universelle.— Démonstration du fait que l'algèbre $\mathfrak{l}\langle X\rangle$ est libre. — Théorème de Poincaré-Birkhoff-Witt.— Produits tensoriels d'espaces vectoriels et d'algèbres.— Algèbres de Hopf.

Pour prouver l'assertion B de la leçon précédente, assertion qui est purement algébrique, nous développerons le formalisme correspondant sous une forme, à vrai dire, plus générale et plus détaillée que nécessaire, ce qui nous permettra sans perte de temps d'exhiber une série de constructions intéressantes et importantes qui sont utiles dans de nombreuses questions de théorie des algèbres de Lie.

Soit $\mathbf{A}$ une catégorie d'algèbres (sur un corps $\mathbb{K}$ qui sera supposé arbitraire dans cette leçon). Conformément à la notion générale d'« objet libre » (notion qui ne sera entièrement définie qu'à l'aide de celle de foncteur adjoint), une algèbre $\mathcal{F}$ de la catégorie $\mathbf{A}$ dans laquelle on a distingué un sous-ensemble X s'appelle *algèbre libre* de la catégorie $\mathbf{A}$ admettant X pour *ensemble de générateurs libres* si pour toute algèbre $\mathcal{A}$ de la catégorie $\mathbf{A}$, toute application $X \to \mathcal{A}$ se prolonge de façon unique en un homomorphisme $\mathcal{F} \to \mathcal{A}$.

Par exemple, l'*algèbre des polynômes* $\mathbb{K}[X]$ de variables parcourant l'ensemble X est une algèbre libre dans la catégorie des algèbres associatives, commutatives et unitaires. De façon analogue, l'*algèbre des polynômes de variables non permutables de* X sera une algèbre libre dans la catégorie ALG_0-ASS des algèbres unitaires associatives (mais généralement non commutatives). Nous désignerons cette algèbre par le symbole $\mathbb{K}\langle X\rangle$, et si l'ensemble $X = \{x_1, \ldots, x_n\}$ est fini, par le symbole $\mathbb{K}\langle x_1, \ldots, x_n\rangle$.

Pour le seul cas qui nous intéresse, cas où l'ensemble X est composé de deux éléments x et y, tout élément de l'algèbre $\mathbb{K}\langle x, y\rangle$ se représente de façon unique par une combinaison linéaire à coefficients dans $\mathbb{K}$ d'expressions de la forme

(1) $$x^{p_1}y^{q_1}\dots x^{p_k}y^{q_k}.$$

appelées *monômes en x et en y*. Ici $p_1, q_1, \dots, p_k, q_k$ sont des entiers $\geqslant 0$ qui sont tous non nuls à l'exception éventuellement des « extrêmes » p_1 et q_k. Si $p_1 = 0$, le terme x^{p_1} n'est pas écrit, si $q_k = 0$, c'est le terme y^{q_k} qui ne l'est pas. Le nombre $n = p_1 + q_1 + \dots$ $\dots + p_k + q_k$ s'appelle *degré* (ou *puissance*) du monôme (1).

On admet qu'un *monôme vide* ($k = 0$) s'identifie à l'unité 1 du corps $\mathcal{K}$. Un tel monôme est de degré 0.

Le fait que les monômes (1) forment une base dans l'algèbre $\mathcal{K}\langle x, y\rangle$ définit de façon unique l'addition et la multiplication par des nombres de $\mathcal{K}$ dans $\mathcal{K}\langle x, y\rangle$. Quant à la multiplication, il suffit en raison de la distributivité de la définir uniquement pour les monômes (1). Si des monômes $x^{p_1}y^{q_1}\dots x^{p_k}y^{q_k}$ et $x^{r_1}y^{s_1}\dots x^{r_l}y^{s_l}$ sont tels que $q_k \neq 0$ et $r_1 \neq 0$ ou, au contraire, que $q_k = 0$ et $r_1 =$ $= 0$, alors on convient que leur produit est le monôme

(2) $$x^{p_1}y^{q_1}\dots x^{p_k}y^{q_k}x^{r_1}y^{s_1}\dots x^{r_l}y^{s_l},$$

obtenu en écrivant le deuxième monôme à la suite du premier (pour $q_k = 0$ et $r_1 = 0$, les termes y^{q_k} et x^{r_1} ne sont naturellement pas écrits). Si l'une seulement des puissances q_k et r_1 est nulle, pour produit on prend le monôme obtenu de l'expression (2) (qui dans ce cas n'est plus un monôme) pour $q_k = 0$ par la substitution de $x^{p_k+r_1}$ à $x^{p_k}y^{q_k}x^{r_1}$ et pour $r_1 = 0$, de $y^{q_k+s_1}$ à $y^{q_k}x^{r_1}y^{s_1}$. Une vérification immédiate montre que cette multiplication est associative, l'élément neutre étant le monôme vide (1).

Si maintenant $\{x, y\} \to \mathcal{A}$ est une application quelconque de l'ensemble $\{x, y\}$ dans une algèbre unitaire associative $\mathcal{A}$, alors en associant à tout monôme (1) un élément $a^{p_1}b^{q_1}\dots a^{p_k}b^{q_k}$ de l'algèbre $\mathcal{A}$, où a et b sont les images des générateurs x et y dans $\mathcal{A}$, et en prolongeant linéairement cette application à des monômes arbitraires, on obtient, comme le montre une vérification immédiate, un homomorphisme $\mathcal{K}\langle x, y\rangle \to \mathcal{A}$ faisant correspondre à x et y les éléments a et b, c'est-à-dire un homomorphisme qui prolonge l'application $\{x, y\} \to \mathcal{A}$. Cet homomorphisme est unique, puisque tout homomorphisme qui prolonge cette application doit associer à un monôme (1) un élément $a^{p_1}b^{q_1}\dots a^{p_k}b^{q_k}$. Ceci prouve que *l'algèbre $\mathcal{K}\langle x, y\rangle$ est une algèbre libre de la catégorie* ALG$_0$-ASS *à générateurs libres x et y*.

Pour tout X, l'algèbre $\mathcal{K}\langle X\rangle$ se décrit de façon analogue. On prouve qu'elle est libre comme pour $X = \{x, y\}$ *mutatis mutandis*.

Les algèbres libres de la catégorie ALG-LIE (qui sont appelées *algèbres de Lie libres*) sont bien plus difficiles à construire. La voie

la plus simple consiste apparemment à considérer dans l'algèbre de Lie $[\mathbb{K}\langle X\rangle]$ adjointe à l'algèbre des polynômes $\mathbb{K}\langle X\rangle$ la sous-algèbre $\mathfrak{L}\langle X\rangle$ engendrée par l'ensemble X, c'est-à-dire la plus petite sous-algèbre contenant cet ensemble. Tout comme dans le cas où $X = \{x, y\}$ (cf. leçon précédente), les éléments de l'algèbre de Lie $\mathfrak{L}\langle X\rangle$ sont les polynômes de Lie en les éléments de X, c'est-à-dire toutes les expressions obtenues en additionnant ces éléments, en les multipliant par des nombres et en leur appliquant le crochet de Lie $[a, b] = ab - ba$ (tous ces éléments appartiennent évidemment à $\mathfrak{L}\langle X\rangle$ et forment une sous-algèbre de l'algèbre $[\mathbb{K}\langle X\rangle]$). L'étude de l'algèbre $\mathfrak{L}\langle X\rangle$ est assez compliquée du fait que malheureusement nous ne pouvons représenter ses éléments par des expressions canoniques simples (à l'image des combinaisons linéaires de monômes pour les éléments de l'algèbre $\mathbb{K}\langle X\rangle$).

Comme à la leçon précédente, on désignera par ι l'injection $\mathfrak{L}\langle X\rangle \to \mathbb{K}\langle X\rangle$, c'est-à-dire, plus exactement, l'application qui à un polynôme de Lie $u \in \mathfrak{L}\langle X\rangle$ associe le polynôme $\iota u \in \mathbb{K}\langle X\rangle$ obtenu en développant tous les crochets de Lie.

Proposition 1. *L'algèbre $\mathfrak{L}\langle X\rangle$ est une algèbre de Lie libre dont X est un ensemble de générateurs libres.*

La démonstration de cette proposition est très compliquée. Nous la ferons par étapes, en commençant par prouver un lemme se rapportant à des algèbres de Lie quelconques.

Soient $\mathfrak{g}$ une algèbre de Lie (sur un corps $\mathbb{K}$) et

$$X = \{x_i,\ i \in I\} \tag{3}$$

une base de $\mathfrak{g}$ (considérée comme un espace vectoriel). Nous n'admettons pas que l'algèbre $\mathfrak{g}$ est de dimension finie, donc l'ensemble I des indices, ensemble qui est supposé totalement ordonné, est infini. (L'existence d'une base dans un espace vectoriel arbitraire se prouve sans peine à l'aide du lemme de Zorn, puisque l'ensemble de tous les sous-espaces de cet espace munis d'une base est inductif, et, de plus, tout sous-espace propre muni d'une base peut être inclus — par adjonction d'un vecteur — dans un espace de dimension plus élevée muni d'une base.)

Soit, par ailleurs, $\mathcal{T}_I$ un espace vectoriel muni d'une base dont les éléments z_α sont indexés par toutes les suites $\alpha = (i_1, i_2, \ldots, i_n)$ monotones (c'est-à-dire telles que $i_1 \leqslant i_2 \leqslant \ldots \leqslant i_n$), où $i_1, i_2, \ldots, i_n \in I$.

On désignera par $|\alpha|$ le nombre n de termes de toute suite $\alpha = (i_1, \ldots, i_n)$. On admettra que la suite vide $\varnothing$, pour laquelle $|\varnothing| = 0$, est monotone.

On notera $[a, b]$ le produit des éléments a, b d'une algèbre de Lie $\mathfrak{g}$.

Lemme 1. *A tout élément $a \in \mathfrak{g}$ et tout élément $v \in \mathcal{T}'_I$, on peut associer un élément $av \in \mathcal{T}'_I$ tel que*

a) *av dépende linéairement de a et de v;*

b) *pour tous éléments a, $b \in \mathfrak{g}$ et tout élément $v \in \mathcal{T}'_I$, l'on ait*

$$[a, b]\, v = a\,(bv) - b\,(av);$$

c) *si $\alpha = (i_1, \ldots, i_n)$ et $i < i_1$, alors*

$$x_i z_\alpha = z_{i\alpha},$$

où $i\alpha = (i, i_1, \ldots, i_n)$.

Démonstration. En vertu de a), il suffit de construire des éléments de la forme $x_i z_\alpha$. Nous le ferons par récurrence sur $|\alpha|$ et sur i. Pour réaliser cette récurrence, il est commode de renforcer le lemme en exigeant que soit remplie la condition subsidiaire suivante:

d) *si* $x_i z_\alpha = \sum c_k z_{\beta_k}$, *alors* $|\beta_k| \leqslant |\alpha| + 1$ *pour tous les* k.

Pour $\alpha = \varnothing$ et pour tout i, on pose par définition

$$x_i z_\varnothing = z_i.$$

Supposons que les éléments $x_j z_\beta$ ont déjà été construits pour tous les j et tous les β tels que $|\beta| < |\alpha|$ ainsi que pour tous les $j < i$ lorsque $|\beta| = |\alpha|$. Posons $\alpha = i_1\beta$ et définissons l'élément $x_i z_\alpha$ à l'aide de la formule

$$x_i z_\alpha = \begin{cases} z_{i\alpha} \text{ pour } i \leqslant i_1; \\ x_{i_1}(x_i z_\beta) + [x_i, x_{i_1}]\, z_\beta \text{ pour } i > i_1. \end{cases}$$

Cette définition nous conduit à un seul élément $x_i z_\alpha$ en vertu de la condition d) et de l'hypothèse de la récurrence. Il est clair que les conditions c) et d) sont remplies pour l'élément $x_i z_\alpha$.

Donc, tous les éléments $x_i z_\alpha$ et partant (par linéarité) tous les éléments av, $a \in \mathfrak{g}$, $v \in \mathcal{T}'_I$, sont construits. Reste à vérifier la condition b). Il est clair qu'il suffit de le faire uniquement pour $a = x_i$, $b = x_j$, $v = z_\alpha$. Donc, il nous reste à prouver que pour tous i, j et toute suite monotone $\alpha = (i_1, \ldots, i_n)$, on a

$$x_i\,(x_j z_\alpha) - x_j\,(x_i z_\alpha) = [x_i, x_j]\, z_\alpha. \tag{4}$$

Cette égalité est évidente pour $i = j$. Par ailleurs, si elle est remplie pour un couple (i, j), elle le sera pour le couple (j, i). Donc, sans perdre en généralité, on peut admettre que $i > j$.

Effectuons une récurrence sur $|\alpha|$. Pour $\alpha = \varnothing$ (et $i > j$), on a par définition

$$x_i\,(x_j z_\varnothing) = x_j\,(x_i z_\varnothing) + [x_i, x_j]\, z_\varnothing.$$

Donc, la relation (4) est réalisée pour $\alpha = \varnothing$.

Supposons maintenant que l'égalité (4) est remplie pour tous les α tels que $|\alpha| < n$ et considérons une suite arbitraire $\alpha = (i_1, \ldots, i_n)$ telle que $|\alpha| = n$. Pour des raisons de symétrie on posera $i_1 = k$ et $\beta = (i_2, \ldots, i_n)$.

Si $j \leqslant k$, on a par définition

$$x_i (x_j z_\alpha) = x_i z_{j\alpha} = x_j (x_i z_\alpha) + [x_i, x_j] z_\alpha,$$

si bien que (4) est valable dans ce cas aussi.

Supposons maintenant que $j > k$. Par hypothèse de la récurrence, l'égalité (4) est valable pour $\alpha = \beta$ (et pour tous i et j). Donc

$$\begin{aligned}[x_i, x_j] z_\alpha &= [x_i, x_j] (x_k z_\beta) = \\ &= x_k([x_i, x_j] z_\beta) + [[x_i, x_j], x_k] z_\beta = \\ &= x_k (x_i (x_j z_\beta)) - x_k (x_j (x_i z_\beta)) + [[x_i, x_j], x_k] z_\beta.\end{aligned}$$

et, par suite, l'égalité (4) peut être mise sous la forme suivante:

$$x_i (x_j (x_k z_\beta)) - x_j (x_i (x_k z_\beta)) + x_k (x_j (x_i z_\alpha)) - \\ - x_k (x_i (x_j z_\alpha)) = [[x_i, x_j], x_k] z_\beta.$$

Désignons cette égalité par le symbole (i, j, k).

A noter que l'on peut considérer que les égalités (j, k, i) et (k, i, j), déduites de (i, j, k) par une permutation cyclique des indices, sont prouvées. En effet, si l'on change le nom des indices $j \mapsto i$, $k \mapsto j$, $i \mapsto k$, l'égalité (j, k, i) se transforme en l'égalité (i, j, k), avec $i > j$ et $j < k$, qui a déjà été prouvée. De façon analogue, si $k \mapsto j$, $i \mapsto i$ et $j \mapsto k$, l'égalité (k, i, j) se transforme (avec inversion de tous les signes) en l'égalité (i, j, k) avec $i > j$ et $j < k$.

Par ailleurs, en additionnant les égalités (j, k, i) et (k, i, j), on obtient de toute évidence une égalité dont le premier membre est justement égal à celui de l'égalité (i, j, k) pris avec le signe opposé et le second membre égal à

$$([[x_j, x_k], x_i] + [[x_k, x_i], x_j]) z_\beta,$$

ce qui en vertu de l'identité de Jacobi est égal au second membre de l'égalité (i, j, k) pris avec le signe opposé. Or, les égalités (j, k, i) et (k, i, j) sont valables, donc il en est de même de l'égalité (i, j, k).

Ce qui achève la démonstration du lemme 1. □

On rappelle qu'un espace vectoriel $\mathcal{V}$ est un *module sur une algèbre associative* $\mathcal{A}$ (ou simplement un *$\mathcal{A}$-module*) si pour tout élément $a \in \mathcal{A}$ et tout élément $v \in \mathcal{V}$ est défini un élément $av \in \mathcal{V}$ dépendant linéairement de a et de v et si pour tous éléments $a, b \in \mathcal{A}$ on a

$$(ab) v = a (bv). \tag{5}$$

Le fait que l'élément av dépend linéairement de v exprime que l'application $v \mapsto av$ de l'espace $\mathcal{V}$ dans lui-même est linéaire. En désignant cette application par $\theta(a)$, on obtient, par conséquent, une application θ de l'algèbre $\mathcal{A}$ dans l'algèbre End $\mathcal{V}$ des applications linéaires (endomorphismes) de $\mathcal{V}$ dans $\mathcal{V}$. Le fait que l'élément av dépend linéairement de l'élément a exprime que l'application θ est linéaire et la relation (5) dit que cette application est un homomorphisme d'algèbres. Réciproquement, la donnée d'un homomorphisme quelconque d'algèbres $\theta: \mathcal{A} \to \text{End}\ \mathcal{V}$ définit sur $\mathcal{V}$ une structure de module sur $\mathcal{A}$ pour laquelle $av = \theta(a)v$.

De façon analogue, on dit qu'un espace vectoriel $\mathcal{V}$ est un *module sur une algèbre de Lie* $\mathfrak{g}$ si est donné un homomorphisme $\theta: \mathfrak{g} \to [\text{End}\ \mathcal{V}]$ de l'algèbre de Lie $\mathfrak{g}$ dans l'algèbre de Lie $[\text{End}\ \mathcal{V}]$. En termes d'éléments, cela signifie que pour tout élément $a \in \mathfrak{g}$ et tout élément $v \in \mathcal{V}$ est défini un élément $av = \theta(a)v$ dépendant linéairement de a et de v, et, de plus, pour tous éléments $a, b \in \mathfrak{g}$ on a

$$[a, b] = a(bv) - b(av).$$

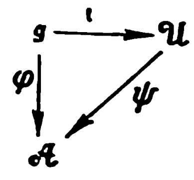
En comparant ces conditions avec les conditions a) et b) du lemme 1, on remarque que ce lemme exprime que *tout espace vectoriel* $\mathcal{V}_I$ *peut être muni d'une structure de module sur une algèbre de Lie* $\mathfrak{g}$ *telle que pour toute suite monotone* $\alpha = (i_1, i_2, \ldots, i_n)$ *et tout* $i \leqslant i_1$ *l'on ait*

$$x_i z_\alpha = z_{i\alpha}.$$

C'est sous cette forme que nous utiliserons ce lemme.

Soient maintenant $\mathfrak{U}$ une algèbre unitaire associative, $\iota: \mathfrak{g} \to [\mathfrak{U}]$ un homomorphisme d'une algèbre de Lie $\mathfrak{g}$ dans l'algèbre de Lie des commutateurs de l'algèbre $\mathfrak{U}$.

Définition 1. On dit qu'un homomorphisme $\iota: \mathfrak{g} \to [\mathfrak{U}]$ est *universel* et l'algèbre $\mathfrak{U}$ (considérée avec cet homomorphisme), une *algèbre enveloppante universelle* de l'algèbre de Lie $\mathfrak{g}$ si pour toute algèbre unitaire associative $\mathcal{A}$ et tout homomorphisme $\varphi: \mathfrak{g} \to [\mathcal{A}]$ il existe un seul homomorphisme $\psi: \mathfrak{U} \to \mathcal{A}$ pour lequel le diagramme



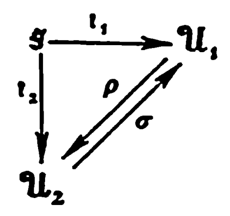
est commutatif.

On écrira $\psi = \mathfrak{U}\varphi$.

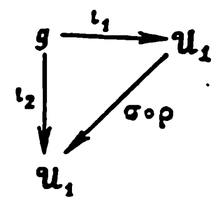
On remarquera que si $\mathcal{A} = \mathfrak{U}$ et $\varphi = \iota$, alors $\mathfrak{U}\varphi$ est l'homomorphisme identique $\text{id}: \mathfrak{U} \to \mathfrak{U}$.

Une algèbre enveloppante universelle de l'algèbre de Lie $\mathfrak{l}\langle X\rangle$ *est l'algèbre* $K\langle X\rangle$ (relativement à l'injection ι: $\mathfrak{l}\langle X\rangle \to [K\langle X\rangle]$). En effet, soit φ: $\mathfrak{l}\langle X\rangle \to [\mathcal{A}]$ un homomorphisme arbitraire de l'algèbre $\mathfrak{l}\langle X\rangle$ dans l'algèbre des commutateurs $[\mathcal{A}]$ d'une algèbre unitaire associative $\mathcal{A}$. L'algèbre $K\langle X\rangle$ étant libre, l'application $\varphi|_X: X \to \mathcal{A}$ se prolonge de façon unique en un homomorphisme $\psi: K\langle X\rangle \to \mathcal{A}$. Il nous faut donc montrer seulement que $\psi\circ\iota = \varphi$, c'est-à-dire que $\psi|_{\mathfrak{l}\langle X\rangle} = \varphi$. Mais par construction $\psi = \varphi$ sur X. De plus, si $\varphi a = \psi a$ et $\varphi b = \psi b$, où a, $b \in \mathfrak{l}\langle X\rangle$, alors $\varphi(a+b) = \psi(a+b)$, $\varphi(ka) = \psi(kb)$ (car les applications φ et ψ sont toutes deux linéaires) et $\varphi[a, b] = [\varphi a, \varphi b] = \varphi a\cdot\varphi b - \varphi b\cdot\varphi a = \psi a\cdot\psi b - \psi b\cdot\psi a = \psi(ab - ba) = \psi[a, b]$. Donc, $\varphi = \psi$ sur tous les polynômes de Lie en X, c'est-à-dire sur l'algèbre $\mathfrak{l}\langle X\rangle$ tout entière. □

Si $\iota_1: \mathfrak{g} \to [\mathfrak{U}_1]$ et $\iota_2: \mathfrak{g} \to [\mathfrak{U}_2]$ sont deux homomorphismes universels, il existe des homomorphismes $\rho = \mathfrak{U}_1\iota_2: \mathfrak{U}_1 \to \mathfrak{U}_2$ et $\sigma = \mathfrak{U}_2\iota_1: \mathfrak{U}_2 \to \mathfrak{U}_1$ tels que le diagramme



est commutatif. Ceci étant, le composé $\sigma\circ\rho$ est justiciable du diagramme commutatif



qui dit que $\sigma\circ\rho = \mathfrak{U}_1\iota_1$. Mais, comme déjà remarqué plus haut, $\mathfrak{U}_1\iota_1 = \mathrm{id}$, où id est l'application identique $\mathfrak{U}_1 \to \mathfrak{U}_1$. Donc $\sigma\circ\rho = \mathrm{id}$. On montre de façon analogue que le composé $\rho\circ\sigma$ est l'application identique $\mathfrak{U}_2 \to \mathfrak{U}_2$. Par suite, les homomorphismes ρ et σ sont réciproques l'un de l'autre. Ceci prouve que pour *tout couple d'algèbres enveloppantes universelles* $\mathfrak{U}_1$ *et* $\mathfrak{U}_2$ *d'une algèbre de Lie* $\mathfrak{g}$, *il existe un isomorphisme* $\rho: \mathfrak{U}_1 \to \mathfrak{U}_2$ tel que $\rho\circ\iota_1 = \iota_2$.

De ce point de vue l'algèbre enveloppante universelle $\mathfrak{U}$ d'une algèbre de Lie est unique.

Pour prouver son existence, choisissons une base arbitraire (3) dans $\mathfrak{g}$ et considérons l'algèbre des polynômes non commutatifs $K\langle X\rangle$. L'algèbre $\mathfrak{g}$ (traitée comme un espace vectoriel) s'identifie

canoniquement au sous-espace $K_1\langle X\rangle$ de l'algèbre $K\langle X\rangle$ composé des polynômes homogènes du premier degré. Donc, pour tous éléments x, $y \in \mathfrak{g}$, dans l'algèbre $K\langle X\rangle$ seront définis trois éléments (visiblement distincts pour $x \neq 0$ et $y \neq 0$): l'élément $[x, y] \in K_1\langle X\rangle$ qui est un polynôme homogène du premier degré, et les éléments xy et yx qui sont des polynômes homogènes du second degré. Soient $\mathfrak{J}$ l'idéal de l'algèbre $K\langle X\rangle$, engendré par les polynômes de la forme $xy - yx - [x, y]$, $\mathfrak{U} = K\langle X\rangle/\mathfrak{J}$ l'algèbre quotient correspondante, $\iota: \mathfrak{g} \to \mathfrak{U}$ la restriction à $\mathfrak{g} = K_1\langle X\rangle$ de l'épimorphisme canonique $K\langle X\rangle \to \mathfrak{U}$. Alors, pour tous éléments x, $y \in \mathfrak{g}$, on a l'égalité suivante dans $\mathfrak{U}$

$$\iota[x, y] = \iota(xy - yx) = \iota x \cdot \iota y - \iota y \cdot \iota x = [\iota x, \iota y],$$

égalité qui exprime que ι est un homomorphisme d'algèbres de Lie $\mathfrak{g} \to [\mathfrak{U}]$.

Supposons maintenant que $\mathcal{A}$ est une algèbre unitaire associative quelconque et soit $\varphi: \mathfrak{g} \to [\mathcal{A}]$ un homomorphisme. La restriction de φ à X se prolonge de façon unique en un homomorphisme $\overline{\psi}: K\langle X\rangle \to \mathcal{A}$. Les applications φ et $\overline{\psi}$ sont linéaires et confondues sur la base X de l'espace $\mathfrak{g} = K_1\langle X\rangle$. Donc, $\varphi = \overline{\psi}$ sur $\mathfrak{g}$, et, par suite, pour tous éléments x, $y \in \mathfrak{g}$, on a

$$\begin{aligned}\overline{\psi}(xy - yx - [x, y]) &= \overline{\psi}x \cdot \overline{\psi}y - \overline{\psi}y \cdot \overline{\psi}x - \overline{\psi}[x, y] = \\ &= \varphi x \cdot \varphi y - \varphi y \cdot \varphi x - \varphi [x, y] = \\ &= [\varphi x, \varphi y] - \varphi [x, y] = 0.\end{aligned}$$

Donc, $\overline{\psi}(\mathfrak{J}) = 0$, et, par suite, $\overline{\psi}$ induit un homomorphisme $\psi: \mathfrak{U} \to \mathcal{A}$ tel que visiblement $\varphi \circ \iota = \psi$. Ceci exprime que l'homomorphisme ι est universel.

Nous avons ainsi démontré que *pour toute algèbre de Lie $\mathfrak{g}$, il existe une algèbre enveloppante universelle $\mathfrak{U}$.*

Remarque 1. Il est aisé de voir que l'algèbre $\mathfrak{U}$ ne dépend pas du choix de la base (3) de l'algèbre de Lie $\mathfrak{g}$ et, par suite, la correspondance $\mathfrak{g} \mapsto \mathfrak{U}$ est un foncteur ALG-LIE $\to$ ALG$_0$-ASS. Ce foncteur est tel que pour toute algèbre associative $\mathcal{A}$, l'ensemble $\mathrm{Hom}(\mathfrak{g}, [\mathcal{A}])$ des homomorphismes $\mathfrak{g} \to [\mathcal{A}]$ est en correspondance biunivoque avec l'ensemble $\mathrm{Hom}(\mathfrak{U}, \mathcal{A})$ des homomorphismes $\mathfrak{U} \to \mathcal{A}$:

$$\mathrm{Hom}(\mathfrak{g}, [\mathcal{A}]) \approx \mathrm{Hom}(\mathfrak{U}, \mathcal{A}).$$

Dans le langage de la théorie des catégories, ceci exprime que le foncteur $\mathfrak{g} \mapsto \mathfrak{U}$ est *adjoint à gauche* du commutateur $\mathcal{A} \mapsto [\mathcal{A}]$.

Appliquée à l'algèbre $\mathcal{A} = \text{End}\ \mathcal{V}$, l'universalité de l'homomorphisme ι: $\mathfrak{g} \to [\mathcal{U}]$ exprime que *tout module $\mathcal{V}$ sur l'algèbre de Lie $\mathfrak{g}$ possède une structure unique de $\mathcal{U}$-module qui prolonge sa structure de $\mathfrak{g}$-module, c'est-à-dire une structure telle que*

$$xv = (\iota x)\, v$$

pour tous éléments $x \in \mathfrak{g}$ et $v \in \mathcal{V}$. En effet, la structure de $\mathfrak{g}$-module sur $\mathcal{V}$ est définie par l'homomorphisme $\mathfrak{g} \to [\text{End}\ \mathcal{V}]$ et la structure de $\mathcal{U}$-module, par l'homomorphisme $\mathcal{U} \to \text{End}\ \mathcal{V}$. □

Grâce au lemme 1, on déduit de là que *sur l'espace vectoriel $\mathcal{V}_I$, il existe une structure de $\mathcal{U}$-module telle que*

(6) $$(\iota x_i)\, z_\alpha = z_{i\alpha}$$

pour toute suite monotone $\alpha = (i_1, \ldots, i_n)$ et tout $i \leqslant i_1$.

En particulier,

$$(\iota x_i)\, z_\varnothing = z_i.$$

La dernière formule entraîne immédiatement la proposition suivante:

Proposition 2. *Pour toute algèbre de Lie $\mathfrak{g}$, l'application $\iota: \mathfrak{g} \to [\mathcal{U}]$ est injective, de sorte que l'algèbre $\mathfrak{g}$ peut être identifiée à une sous-algèbre de l'algèbre des commutateurs $[\mathcal{U}]$.*

Démonstration. Soient x un élément de l'algèbre de Lie $\mathfrak{g}$ tel que $\iota x = 0$, $x = \sum c_i x_i$, sa décomposition par rapport à la base (3). Alors

$$\sum c_i z_i = \sum c_i\, (\iota x_i) z_\varnothing = (\iota \sum c_i x_i)\, z_\varnothing = (\iota x)\, z_\varnothing = 0 z_\varnothing = 0,$$

ce qui est possible (puisque les éléments z_i forment une partie de la base de $\mathcal{V}_I$) si seulement $c_i = 0$ pour tous les i. Donc $x = 0$. □

La proposition 2 explique l'épithète « enveloppante » attribuée à l'algèbre $\mathcal{U}$.

Nous admettrons dans la suite que l'application ι est une injection et nous désignerons, en particulier, tout élément ιx, $x \in \mathfrak{g}$, simplement par x.

La proposition 2 entraîne aussitôt la proposition 1.

Démonstration de la proposition 1. Soient φ: $X \to \mathfrak{g}$ une application quelconque de X dans une algèbre de Lie $\mathfrak{g}$, $\iota: \mathfrak{g} \to \mathcal{U}$ l'injection de l'algèbre $\mathfrak{g}$ dans son algèbre enveloppante universelle $\mathcal{U}$. Puisque l'algèbre $K\langle X \rangle$ est une algèbre libre de la catégorie ALG_0-ASS, il existe un homomorphisme $\bar{\psi}$: $K\langle X \rangle \to \mathcal{U}$ confondu sur X avec la composée $\iota \circ \varphi$: $X \to \mathcal{U}$. Cet homomorphisme fait correspondre à tout élément de l'algèbre $\mathfrak{l}\langle X \rangle$, c'est-à-dire à tout polynôme de Lie en $x \in X$, un polynôme

de Lie en les éléments correspondants $(\iota \circ \varphi)\, x$ de la sous-algèbre $\iota(\mathfrak{g})$, donc un élément de cette sous-algèbre. Ceci exprime que $\overline{\psi}(\mathfrak{l}\langle X\rangle) \subset \iota(\mathfrak{g})$. Comme l'application $\iota : \mathfrak{g} \to \iota(\mathfrak{g})$ est bijective, il s'ensuit de là que l'application $\overline{\psi}$ induit une application $\psi : \mathfrak{l}\langle X\rangle \to \mathfrak{g}$ telle que $\iota \circ \psi = \overline{\psi}$ sur $\mathfrak{l}\langle X\rangle$. Pour achever la démonstration, il reste à remarquer que l'application ψ est visiblement un homomorphisme d'algèbres de Lie et est confondue avec φ sur X. □

La proposition 2 revient à affirmer que les éléments $x_i = \iota x_i$ de l'algèbre $\mathfrak{U}$ sont linéairement indépendants. Sous cette forme, elle admet une importante généralisation qui est fondamentale dans la démonstration de l'assertion B de la leçon précédente (signalons à propos que nous n'avons pas besoin de la proposition 1 pour prouver l'assertion B ; nous l'avons établie pour son intérêt intrinsèque). Pour formuler cette généralisation, considérons pour toute suite monotone $\alpha = (i_1, i_2, \ldots, i_n)$ d'éléments de I l'élément

$$x_\alpha = x_{i_1} x_{i_2} \ldots x_{i_n} \tag{7}$$

de l'algèbre enveloppante universelle $\mathfrak{U}$ (pour $\alpha = \varnothing$ on admet que $x_\varnothing = 1$). Pour faciliter les formulations ultérieures, nous appellerons les éléments de la forme (7) *éléments spéciaux* de l'algèbre $\mathfrak{U}$ (relativement à la base (3) de l'algèbre $\mathfrak{g}$).

Proposition 3. *Les éléments spéciaux x_α de l'algèbre enveloppante universelle $\mathfrak{U}$ sont linéairement indépendants.*

D é m o n s t r a t i o n. De la formule (6) on déduit par une récurrence immédiate que

$$x_\alpha z_\varnothing = z_\alpha$$

pour toute suite monotone α. Il reste maintenant à reproduire des raisonnements déjà connus: si $\sum c_\alpha x_\alpha = 0$, alors

$$\sum c_\alpha z_\alpha = \sum c_\alpha x_\alpha z_\varnothing = 0 z_\varnothing = 0,$$

ce qui n'est possible que si $c_\alpha = 0$ pour tout α. □

La proposition 3 peut être précisée:

Proposition 4. *Les éléments spéciaux x_α forment une base dans l'algèbre $\mathfrak{U}$ (traitée comme un espace vectoriel).*

Faisons quelques remarques préliminaires avant de prouver cette proposition.

Les éléments x_α sont définis de toute évidence pour t o u t e s les suites (pas forcément monotones) $\alpha = (i_1, \ldots, i_n)$ d'éléments de I (ces éléments seront dits spéciaux uniquement pour α monotones).

Pour toute suite $\alpha = (i_1, \ldots, i_n)$ d'éléments de I, on désignera par $d(\alpha)$ le nombre de *désordres* dans α, c'est-à-dire le nombre de couples (k, l), $1 \leqslant k$, $l \leqslant n$, tels que $k < l$ et $i_k > i_l$. Une suite α est monotone si et seulement si $d(\alpha) = 0$. En particulier, $d(\varnothing) = 0$.

Une récurrence peu compliquée sur $|\alpha|$ et $d(\alpha)$ montre maintenant que *tout élément de la forme x_α est une combinaison linéaire d'éléments spéciaux.*

En effet, cette proposition est automatiquement vraie pour $d(\alpha) = 0$. Supposons qu'elle a déjà été prouvée pour tous les éléments de la forme x_β tels que $|\beta| < |\alpha|$ ou $|\beta| = |\alpha|$ et $d(\beta) < d(\alpha)$, et considérons un élément x_α tel que $d(\alpha) > 0$. Il est clair qu'un couple d'indices voisins (i, j) de la suite α est en désordre, c'est-à-dire que $i > j$. Donc, l'élément x_α peut être écrit sous la forme suivante:

$$x_\alpha = x_{\alpha'} x_i x_j x_{\alpha''},$$

où α' et α'' sont des suites d'indices (éventuellement vides). Or, par définition du commutateur

$$x_i x_j = x_j x_i + [x_i, x_j],$$

donc,

$$x_\alpha = x_{\alpha'} x_j x_i x_{\alpha'} + x_{\alpha'} [x_i, x_j] x_{\alpha'}.$$

Dans $\mathfrak{g}$, l'élément $[x_i, x_j]$ se développe par rapport à la base (3), c'est-à-dire que

$$[x_i, x_j] = c_{k_1} x_{k_1} + c_{k_2} x_{k_2} + \ldots,$$

où seuls un nombre fini de coefficients $c_{k_1}, c_{k_2}, \ldots$ sont non nuls. Donc,

$$x_\alpha = x_{\beta_0} + c_{k_1} x_{\beta_1} + c_{k_2} x_{\beta_2} + \ldots,$$

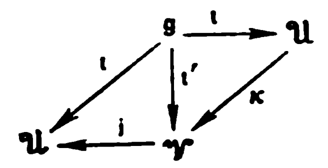
où

$$x_{\beta_0} = x_{\alpha'} x_j x_i x_{\alpha''} \text{ et } x_{\beta_s} = x_{\alpha'} x_{k_s} x_{\alpha''} \text{ pour } s = 1, 2, \ldots$$

Comme $d(\beta_0) = d(\alpha) - 1$ et $|\beta_s| = |\alpha| - 1$ pour $s > 0$, les éléments x_{β_0} et x_{β_s} s'expriment, par hypothèse de la récurrence, en fonction des éléments spéciaux. Donc, il en est de même de l'élément x_α. □

Par ailleurs, de la construction de l'algèbre enveloppante universelle $\mathfrak{U}$, il s'ensuit immédiatement que *l'algèbre $\mathfrak{U}$ est engendrée (en tant qu'algèbre unitaire) par tous les éléments de* $\mathfrak{g}$, plus exactement, par les éléments de la forme ιx, où $x \in \mathfrak{g}$. Du reste, cette affirmation résulte directement de la définition 1. En effet, soient $\mathcal{F}$ la sous-algèbre de $\mathfrak{U}$, engendrée par tous les éléments de $\mathfrak{g}$ (et par l'unité 1), j, l'injection $\mathcal{F} \to \mathfrak{U}$. Comme $\iota\mathfrak{g} \subset \mathcal{F}$, l'homomorphisme ι induit un homomorphisme $\iota' : \mathfrak{g} \to [\mathcal{F}]$ (tel que $j \circ \iota' = \iota$).

Soit $k = \mathfrak{U}\iota'$. Alors, il ressort immédiatement du diagramme commutatif



que le composé $j \circ k$ est l'application $\mathfrak{U}\iota = \mathrm{id}$. Donc, j est un épimorphisme et, par suite, $j' = \mathfrak{U}$. □

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 4. En vertu de la proposition 3, il suffit simplement de montrer que tout élément de l'algèbre $\mathfrak{U}$ est une combinaison linéaire d'éléments spéciaux. Or, puisque les éléments de $\mathfrak{g}$ engendrent l'algèbre $\mathfrak{U}$, tout élément de cette dernière est un polynôme en les éléments x_i de la base (3), c'est-à-dire est une combinaison linéaire de monômes de la forme x_α (pour α quelconque). Ceci prouve la proposition 4. puisque, comme établi plus haut, tout élément de la forme x_α est une combinaison linéaire d'éléments spéciaux. □

La proposition 4 s'appelle généralement *théorème de Poincaré-Birkhoff-Witt*. Signalons que les propositions 2 et 3 portent le même nom.

Nous aurons besoin de la proposition 4 uniquement pour l'appliquer à l'algèbre $\mathfrak{l}\langle X\rangle$ et à son algèbre enveloppante universelle $K\langle X\rangle$ (de sorte qu'en toute rigueur nous aurions pu nous passer de la notion d'algèbre enveloppante; mais ceci n'aurait apporté aucune simplification sensible).

L'étape suivante de la démonstration de l'assertion B consistera à caractériser d'une autre manière, plus effective, les éléments de l'algèbre $\mathfrak{l}\langle X\rangle$. A cet effet nous aurons besoin d'une construction générale.

Soient $\mathcal{A}$ et $\mathcal{B}$ deux espaces vectoriels, admettant respectivement des bases $\{x_i\}$ et $\{y_j\}$. Considérons tous les produits formels de la forme x_iy_j et l'espace linéaire $\mathcal{C}$ ayant $\{x_iy_j\}$ pour base. Tout élément de $\mathcal{C}$ est une combinaison linéaire formelle de la forme

$$\sum k_{ij}x_iy_j, \quad k_{ij} \in K,$$

où un nombre fini seulement de coefficients k_{ij} sont non nuls. Comparons l'espace vectoriel $\mathcal{C}$ à l'espace vectoriel $\mathcal{C}'$ engendré par les sommes formelles de la forme

$$\sum a_jy_j, \tag{8}$$

où $a_j \in \mathcal{A}$. et dans lequel les opérations linéaires se définissent de manière évidente (cet espace vectoriel est la somme directe d'es-

paces vectoriels isomorphes à $\mathcal{A}$, le nombre de termes de cette somme étant égal à celui des éléments de la base $\{y_j\}$). Signalons que l'espace vectoriel $\mathcal{C}'$ *ne dépend pas* du choix de la base $\{x_i\}$ (cette base ne participe pas à sa construction). Si $a_j = \sum k_{ij}x_i$ sont les décompositions des éléments a_j par rapport à la base $\{x_i\}$, nous associons à l'élément (8) de $\mathcal{C}'$ l'élément $\sum k_{ij}x_iy_j$ de $\mathcal{C}$. Il est évident que cette correspondance est un isomorphisme de $\mathcal{C}'$ sur $\mathcal{C}$. En identifiant $\mathcal{C}'$ et $\mathcal{C}$ avec cet isomorphisme, on obtient que $\mathcal{C}$ ne dépend pas non plus du choix de la base $\{x_i\}$. D'autre part, l'égalité

$$\sum k_{ij}x_iy_j = \sum (\sum k_{ij}x_i)\, y_j$$

a un sens (et est vérifiée) avec cette identification. De façon analogue, l'espace $\mathcal{C}$ s'identifie à l'espace vectoriel $\mathcal{C}''$ des éléments de la forme

$$\sum x_ib_i,$$

où $b_i \in \mathcal{B}$. Ceci exprime que l'espace $\mathcal{C}$ ne dépend pas du choix de la base $\{y_j\}$ et confère en même temps un sens à l'égalité

$$\sum k_{ij}x_iy_j = \sum x_i (\sum k_{ij}y_j).$$

On voit donc que l'espace $\mathcal{C}$ ne dépend pas (en vertu des identifications effectuées) du choix des bases dans les espaces $\mathcal{A}$ et $\mathcal{B}$, c'est-à-dire est défini exclusivement par ces espaces. L'espace $\mathcal{C}$ s'appelle *produit tensoriel* des espaces $\mathcal{A}$ et $\mathcal{B}$ et se note $\mathcal{A} \otimes \mathcal{B}$.

Remarque 1. Les éléments x_iy_i de l'espace $\mathcal{A} \otimes \mathcal{B}$ sont ordinairement désignés par $x_i \otimes y_j$ et l'élément $\sum k_{ij}x_iy_j$ par $\sum k_{ij}x_i \otimes y_j$. De façon plus générale, $a = \sum k_ix_i$ et $b = \sum l_jy_j$ étant deux éléments quelconques respectivement de $\mathcal{A}$ et de $\mathcal{B}$, on désignera par $a \otimes b$ l'élément $\sum k_il_jx_i \otimes y_j$. Il est clair que

$$(a_1 + a_2) \otimes b = a_1 \otimes b + a_2 \otimes b,$$

$$a \otimes (b_1 + b_2) = a \otimes b_1 + a \otimes b_2$$

$$k\,(a + b) = (ka) \otimes b_1 = a \otimes (kb)$$

pour tous éléments a, a_1, $a_2 \in \mathcal{A}$, b, b_1, $b_2 \in \mathcal{B}$ et $k \in K$. Réciproquement, il est immédiat de voir que l'espace vectoriel engendré par les symboles $a \otimes b$ satisfaisant aux trois relations ci-dessus est canoniquement isomorphe à l'espace $\mathcal{A} \otimes \mathcal{B}$. On obtient ainsi une construction de l'espace $\mathcal{A} \otimes \mathcal{B}$ qui est intrinsèque (ne dépend pas de la base). Nous ne nous arrêterons pas en détail sur cette construction, car nous n'en aurons pas besoin.

Remarque 2. On obtiendrait une autre caractéristique intrinsèque pour l'espace $\mathcal{A} \otimes \mathcal{B}$ (tout au moins pour le cas où $\mathcal{A}$ et $\mathcal{B}$ sont

de dimension finie) en introduisant l'espace vectoriel $\mathscr{L}$ des fonctionnelles bilinéaires mixtes $x, y \mapsto \boldsymbol{B}(x, y)$ (cf. II, 5), dont le premier argument parcourt l'espace $\mathscr{A}$ et le second, l'espace $\mathscr{B}$. Il s'avère que $\mathscr{A} \otimes \mathscr{B}$ est dual à $\mathscr{L}$, et de plus le couplage correspondant est défini de façon unique par la relation $\langle a \otimes b, \boldsymbol{B} \rangle = \boldsymbol{B}(a, b)$. Nous n'utiliserons pas non plus cette remarque.

Si $\mathscr{A}$ et $\mathscr{B}$ sont des algèbres, l'espace $\mathscr{A} \otimes \mathscr{B}$ est muni d'une multiplication telle que

$$(a \otimes b)(x \otimes y) = (ax) \otimes (by)$$

pour tous $a, x \in \mathscr{A}$ et $b, y \in \mathscr{B}$. L'espace vectoriel $\mathscr{A} \otimes \mathscr{B}$ sera une algèbre pour cette multiplication. Cette algèbre s'appelle *produit tensoriel des algèbres $\mathscr{A}$ et $\mathscr{B}$* et se note $\mathscr{A} \otimes \mathscr{B}$.

Si les algèbres $\mathscr{A}$ et $\mathscr{B}$ sont associatives et unitaires, il en est de même de l'algèbre $\mathscr{A} \otimes \mathscr{B}$. (On remarquera que cette affirmation est mise en défaut pour les algèbres de Lie.) L'unité de l'algèbre $\mathscr{A} \otimes \mathscr{B}$ est de toute évidence l'élément $1 \otimes 1$.

Dans les notations simplifiées (n'utilisant pas le signe $\otimes$) que nous avons adoptées, la multiplication dans $\mathscr{A} \otimes \mathscr{B}$ est définie par la formule

$$\left(\sum k_{ij} x_i y_j\right)\left(\sum l_{pq} x_p y_q\right) = \sum k_{ij} l_{pq} (x_i x_p)(y_j y_q).$$

Dans le cas $\mathscr{A} = K\langle X \rangle$ et $\mathscr{B} = K\langle Y \rangle$, qui seul nous intéresse maintenant, l'algèbre $\mathscr{A} \otimes \mathscr{B}$ n'est autre que l'algèbre des polynômes en les générateurs de X et de Y, tels que chaque générateur de X commute à chaque générateur de Y. Nous désignerons cette algèbre par $K\langle X, Y \rangle$. Donc,

$$K\langle X, Y \rangle = K\langle X \rangle \otimes K\langle Y \rangle$$

quels que soient X et Y.

La multiplication dans une algèbre $\mathscr{A}$ n'est autre qu'une application linéaire

$$\mu : \mathscr{A} \otimes \mathscr{A} \to \mathscr{A}$$

(l'image d'un élément $a \otimes b \in \mathscr{A} \otimes \mathscr{B}$ par cette application est le produit ab). D'après la dualité de la théorie des catégories (qui consiste à inverser le sens de toutes les flèches), toute application linéaire

$$\delta : \mathscr{A} \to \mathscr{A} \otimes \mathscr{A} \tag{9}$$

sera un objet dual. On appelle *coalgèbre* un espace vectoriel $\mathscr{A}$ pour lequel est définie une application (9), et *coproduit*, l'application (9) (on dit aussi *application diagonale*).

Définition 2. Une algèbre unitaire associative $\mathcal{A}$ est une *algèbre de Hopf* si elle est munie d'une structure de coalgèbre (c'est-à-dire si elle est munie du coproduit (9)) et si :

a) le coproduit (9) est un homomorphisme d'algèbres unitaires;

b) dans l'algèbre $\mathcal{A}$ est donné un sous-espace vectoriel $\mathcal{A}^*$ tel que $\mathcal{A} = K1 \oplus \mathcal{A}^*$ et que pour tout élément $x \in \mathcal{A}^*$, on a

(10) $$\delta x = 1 \otimes x + x \otimes 1 + \sum x_i \otimes y_i,$$

où $x_i, y_i \in \mathcal{A}^*$. Si

$$\delta x = 1 \otimes x + x \otimes 1,$$

l'élément $x \in \mathcal{A}^*$ est dit *primitif.*

On remarquera qu'en vertu de a), on a $\delta 1 = 1 \otimes 1$.

Dans le langage de la théorie des coalgèbres, la condition b) exprime que l'élément $1 \in \mathcal{A}$ est la *counité* du coproduit δ.

Les applications $x \mapsto x \otimes 1$ et $x \mapsto 1 \otimes x$ sont des monomorphismes de l'algèbre $\mathcal{A}$ dans l'algèbre $\mathcal{A} \otimes \mathcal{A}$. On désignera l'élément $x \otimes 1$ par x' et l'élément $1 \otimes x$ par x''. La formule (10) devient alors

(11) $$\delta x = x' + x'' + \sum x_i' y_i'',$$

où $x_i, y_i \in \mathcal{A}^*$, et la condition que x est primitif,

$$\delta x = x' + x''.$$

En vertu de la remarque ci-dessus, si $\mathcal{A} = K\langle X \rangle$, alors $\mathcal{A} \otimes \mathcal{A} = K\langle X', X'' \rangle$, où X' et X'' sont deux exemplaires disjoints de X. En particulier, si $\mathcal{A} = K\langle x, y \rangle$ (seul ce cas nous intéresse), $\mathcal{A} \otimes \mathcal{A} = K\langle x', y'; x'', y'' \rangle$, où x' et y' commutent à x'' et y''.

Munissons l'algèbre $K\langle X \rangle$ du coproduit en exigeant que tous les générateurs $x \in X$ soient primitifs, c'est-à-dire que pour tout élément $x \in X$ on ait

$$\delta x = x' + x''.$$

L'algèbre $K\langle X \rangle$ étant libre, cette condition définit de façon unique le coproduit

(12) $$\delta : K\langle X \rangle \to K\langle X', X'' \rangle,$$

qui est un homomorphisme d'algèbres, c'est-à-dire vérifiant la condition a) de la définition d'une algèbre de Hopf.

S'agissant de la condition b), pour sous-espace vectoriel $\mathcal{A}^*$ y figurant, on prendra la sous-algèbre $K\langle X \rangle^*$ de l'algèbre $K\langle X \rangle$, formée de tous les polynômes sans termes constants (qui sont des combinaisons linéaires des monômes $\neq 1$), ou, ce qui visiblement est équivalent, l'enveloppe linéaire de tous les éléments spéciaux x_α (par rapport à une base de l'algèbre $\mathfrak{l}\langle X \rangle$) correspondant à des suites monotones non vides α. Il est clair que la condition b) est

alors remplie (sous la forme (11)). *Nous avons ainsi muni l'algèbre* $K\langle X\rangle$ *d'une structure d'algèbre de Hopf.*

Cette structure peut être conférée, à partir de considérations générales, non seulement à l'algèbre $K\langle X\rangle$, mais aussi à l'algèbre enveloppante universelle $\mathfrak{U}$ de toute algèbre de Lie $\mathfrak{g}$. En effet, les éléments x' et x'' étant permutables dans l'algèbre $\mathfrak{U} \otimes \mathfrak{U}$, l'application

$$x \mapsto x' + x'', \qquad x \in \mathfrak{g},$$

est un homomorphisme de l'algèbre $\mathfrak{g}$ dans l'algèbre des commutateurs $[\mathfrak{U} \otimes \mathfrak{U}]$. Elle se prolonge donc en un homomorphisme d'algèbres

$$\delta : \mathfrak{U} \to \mathfrak{U} \otimes \mathfrak{U}. \tag{13}$$

Il est clair maintenant que si pour $\mathfrak{U}^*$ nous prenons le sous-espace de $\mathfrak{U}$ des combinaisons linéaires des produits des éléments de $\mathfrak{g}$, alors $\mathfrak{U}$ sera une algèbre de Hopf munie du coproduit δ. Donc, *l'algèbre enveloppante universelle* $\mathfrak{U}$ *d'une algèbre de Lie* $\mathfrak{g}$ *a une structure naturelle d'algèbre de Hopf.*

A noter que si $\mathfrak{U} = K\langle X\rangle$, le coproduit (13) est confondu avec le coproduit (12), puisque tous deux agissent de la même façon sur les générateurs de X.

7—01642

LEÇON 6

Théorème de Friedrichs.— Démonstration de l'assertion B de la leçon 4.— Théorème de Dynkine.— Partie linéaire de la série de Campbell-Hausdorff.— Convergence de la série de Campbell-Hausdorff.— Groupalgèbres de Lie.— Equivalence des catégories des groupuscules et des groupalgèbres de Lie.— Isomorphisme des catégories des groupalgèbres et des algèbres de Lie.— Troisième théorème de Lie.

D'après la construction de la fin de la leçon précédente, *tous les éléments de l'algèbre de Lie* $\mathfrak{g}$ *sont éléments primitifs de son algèbre enveloppante universelle* $\mathfrak{U}$. Il se trouve que la réciproque est vraie:

Proposition 1. *Pour toute algèbre de Lie* $\mathfrak{g}$, *tous les éléments primitifs de l'algèbre enveloppante de Hopf* $\mathfrak{U}$ *sont éléments de* $\mathfrak{g}$.

D é m o n s t r a t i o n. Soient comme dans la leçon précédente $X = \{x_i,\ i \in I\}$ une base quelconque de l'algèbre de Lie $\mathfrak{g}$, $\{x_\alpha\}$ une base de l'algèbre enveloppante $\mathfrak{U}$ formée d'éléments spéciaux x_α (relativement à la base X). Calculons l'image δx_α d'un élément spécial x_α par le coproduit

$$\delta : \mathfrak{U} \to \mathfrak{U} \otimes \mathfrak{U}.$$

Si $\alpha = \varnothing$, c'est-à-dire $x_\alpha = 1$, la réponse est immédiate: $\delta 1 = 1$. *Idem* pour $|\alpha| = 1$, c'est-à-dire lorsque α est composé du seul élément $i \in I$: étant un élément de $\mathfrak{g}$, l'élément x_i est primitif, donc $\delta x_i = x_i' + x_i''$. Par conséquent, il nous faut traiter seulement le cas $|\alpha| > 1$.

Certains termes de la suite α peuvent être identiques. Soient $j_1 < j_2 < \ldots < j_m$ les termes distincts de cette suite, $k_s = k_s(\alpha)$ le nombre de termes de α égaux au terme j_s, $s = 1, \ldots, m$. Donc, $k_s \geqslant 1$, $k_1 + \ldots + k_s = |\alpha|$ et

$$x_\alpha = x_{j_1}^{k_1} \ldots x_{j_m}^{k_m}.$$

Comme $\delta x_j = x'_j + x''_j$, il vient

$$\delta x_\alpha = (x'_{j_1} + x''_{j_1})^{k_1} \ldots (x'_{j_m} + x''_{j_m})^{k_m}.$$

En chassant les parenthèses et en n'écrivant que les termes qui soit contiennent des éléments affectés du même nombre d'accents soit sont linéaires en les éléments de la forme x''_i, on obtient (en tenant compte de la permutabilité des éléments affectés d'un nombre d'accents différent) l'expression suivante pour δx_α:

$$\begin{aligned}\delta x_\alpha = {} & x'^{k_1}_{j_1} \ldots x'^{k_m}_{j_m} + x''^{k_1}_{j_1} \ldots x''^{k_m}_{j_m} + \\ & + k_1 x'^{k_1 - 1}_{j_1} x'^{k_2}_{j_2} \ldots x'^{k_m}_{j_m} \cdot x''_{j_1} + \\ & + k_2 x'^{k_1}_{j_1} x'^{k_2 - 1}_{j_2} \ldots x'^{k_m}_{j_m} \cdot x''_{j_2} + \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ & + k_m x'^{k_1}_{j_1} x'^{k_2}_{j_2} \ldots x'^{k_m - 1}_{j_m} \cdot x''_{j_m} + \ldots,\end{aligned}$$

où, en vertu de la condition $|\alpha| > 1$, aucun terme ne peut être simplifié par un autre. En posant

$$x_{\alpha_1} = x^{k_1 - 1}_{j_1} x^{k_2}_{j_2} \ldots x^{k_m}_{j_m}, \quad \ldots, \quad x_{\alpha_m} = x^{k_1}_{j_1} x^{k_2}_{j_2} \ldots x^{k_m - 1}_{j_m},$$

nous pouvons mettre cette formule sous la forme compendieuse suivante:

$$(1)\quad \delta x_\alpha = x'_\alpha + x''_\alpha + k_1(\alpha)\, x'_{\alpha_1} x''_{j_1} + \ldots + k_m(\alpha)\, x'_{\alpha_m} x''_{j_m} + \ldots, \qquad |\alpha| > 1.$$

Considérons maintenant un élément quelconque

$$a = \sum_{\alpha \neq \varnothing} c_\alpha x_\alpha$$

de l'espace vectoriel $\mathfrak{U}^*$.

Posons $a = a_1 + a_2$, où

$$a_1 = \sum_{|\alpha| = 1} c_\alpha x_\alpha \quad \text{et} \quad a_2 = \sum_{|\alpha| > 1} c_\alpha x_\alpha.$$

L'élément a_1 appartient à l'algèbre $\mathfrak{g}$, donc est primitif. Par suite, l'élément a est primitif si et seulement si a_2 l'est. En particulier, si $a_2 = 0$, l'élément a est primitif.

Soient $a_2 \neq 0$ et $j_1 < \ldots < j_m$ les éléments distincts de toutes les suites α telles que x_α figure dans l'élément a_2 avec un coefficient c_α non nul. Pour un tel x_α, la formule (1) reste en vigueur si l'on admet que $k_s(\alpha) = 0$ et $x_{\alpha_s} = 1$ pour tout indice $s = 1, \ldots, m$ pour lequel j_s n'entre pas dans la suite α. Donc,

$$\begin{aligned}\delta a_2 = {} & \sum_{|\alpha|>1} c_\alpha (x'_\alpha + x''_\alpha + k_1(\alpha)\, x'_{\alpha_1} x''_{j_1} + \ldots + k_m(\alpha)\, x'_{\alpha_m} x''_{j_m} + \ldots) = \\ = {} & a'_2 + a''_2 + \Big(\sum_{|\alpha|>1} k_1(\alpha)\, c_\alpha x'_{\alpha_1}\Big) x''_{j_1} + \ldots + \Big(\sum_{|\alpha|>1} k_m(\alpha)\, c_\alpha x'_{\alpha_m}\Big) x''_{j_m} +\end{aligned}$$

et, par suite, si l'élément a_2 est primitif, pour tout $s = 1, \ldots, m$ on a l'égalité suivante dans l'algèbre $\mathfrak{U}$

$$\sum_{|\alpha|>1} k_s(\alpha) c_\alpha x_{\alpha_s} = 0.$$

Dans cette égalité, les termes correspondant à des suites α distinctes ne peuvent se simplifier, puisque $x_{\alpha_s} = x_{\beta_s}$ n'est possible que pour $\alpha_s = \beta_s$, et si $\alpha_s = \beta_s$ et $k_s(\alpha) \neq 0$, $k_s(\beta) \neq 0$, alors $\alpha = \beta$. Donc, si a_2 est primitif, $k_s(\alpha) c_\alpha = 0$ pour toute suite α (participant dans la décomposition de l'élément a_2) et tout s, c'est-à-dire que $k_s(\alpha) = 0$ (puisque nous avons admis que $c_\alpha \neq 0$). Par suite, $|\alpha| = 0$, ce qui contredit la condition $|\alpha| > 1$.

Donc, aucun élément $a_2 \neq 0$ n'est primitif. Par suite, il en est de même de l'élément a.

Ceci prouve qu'un élément $a \in \mathfrak{U}^*$ est primitif si et seulement si $a_2 = 0$, c'est-à-dire si $a = a_1$ et, par suite, $a \in \mathfrak{g}$.

Ceci prouve entièrement la proposition 1. □

Corollaire (théorème de Friedrichs). *Un élément de l'algèbre des polynômes $K\langle X\rangle$ est un polynôme de Lie sur X (plus exactement, est de la forme ιa, où $a \in \mathfrak{l}\langle X\rangle$) si et seulement s'il est primitif.* □

Tout est prêt maintenant pour la démonstration de l'assertion B de la leçon 4. On rappelle que cette assertion se rapporte aux polynômes

$$z_n(x, y) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \sum \frac{x^{p_1}y^{q_1}\ldots x^{p_k}y^{q_k}}{p_1!q_1!\ldots p_k!q_k!} \tag{2}$$

de l'algèbre $K\langle x, y\rangle$, où dans la somme intérieure la sommation est effectuée sur toutes les collections d'entiers positifs $p_1, \ldots, p_k$, $q_1, \ldots, q_k$ telles que $p_1 + \ldots + p_k + q_1 + \ldots + q_k = n$ et $p_i + q_i > 0$ pour tout $i = 1, \ldots, n$ (le corps K est supposé être maintenant de caractéristique 0). Ces polynômes sont les composantes homogènes de degré n de la série formelle $\ln(e^x e^y)$ en les variables non permutables x et y, obtenue en substituant le produit

$$e^x e^y = \sum_{p,q=0}^{\infty} \frac{x^p y^q}{p!q!}$$

des séries

$$e^x = \sum_{p=0}^{\infty} \frac{x^p}{p!} \quad \text{et} \quad e^y = \sum_{q=0}^{\infty} \frac{y^q}{q!}$$

à z dans la série formelle

$$\ln z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (z-1)^k.$$

L'assertion B dit que *tout polynôme* $z_n(x, y)$ *appartient à l'algèbre* $\mathfrak{l}\langle x, y\rangle$, c'est-à-dire, d'après le théorème de Friedrichs, est un élément primitif de l'algèbre de Hopf $K\langle x, y\rangle$ (vérifie la relation $\delta z_n(x, y) = z_n(x', y') + z_n(x'', y'')$).

Conceptuellement, la voie la plus naturelle pour prouver cette assertion consiste à étendre l'algèbre $K\langle x, y\rangle$ à l'algèbre $K\langle\langle x, y\rangle\rangle$ des séries formelles des variables non permutables x et y et à reprendre pour cette algèbre tout ce qui a été fait dans la leçon précédente, c'est-à-dire à introduire la sous-algèbre $\mathfrak{l}\langle\langle x, y\rangle\rangle$ de l'algèbre de Lie des commutateurs $[K\langle\langle x, y\rangle\rangle]$ engendrée par les éléments x et y (cette algèbre est formée des séries dont les composantes homogènes sont toutes les polynômes de Lie de $\mathfrak{l}\langle x, y\rangle$) et à montrer que les séries de $\mathfrak{l}\langle\langle x, y\rangle\rangle$ sont exactement les éléments primitifs de l'algèbre relativement à la comultiplication δ qui est telle que $\delta x = x' + x''$ et $\delta y = y' + y''$. Après cela, l'assertion B (qui revient à affirmer que la série $\ln(e^x e^y)$ appartient à l'algèbre $\mathfrak{l}\langle\langle x, y\rangle\rangle$, c'est-à-dire est un élément primitif de l'algèbre $K\langle\langle x, y\rangle\rangle$), se démontre par un calcul immédiat:

$$\begin{aligned}(3)\qquad \delta \ln(e^x e^y) &= \ln(e^{\delta x} e^{\delta y}) = \ln(e^{x'+x''} e^{y'+y''}) = \\ &= \ln(e^{x'} e^{x''} e^{y'} e^{y''}) = \ln(e^{x'} e^{y'} e^{x''} e^{y''}) = \\ &= \ln(e^{x'} e^{y'}) + \ln(e^{x''} e^{y''}) = \ln(e^x e^y)' + \ln(e^x e^y)''.\end{aligned}$$

(Ce calcul utilise le fait évident que si des éléments ξ et η sont permutables, $e^{\xi+\eta} = e^{\xi} e^{\eta}$ et $\ln(\xi\eta) = \ln\xi + \ln\eta$.)

Mais le passage à l'algèbre $K\langle\langle x, y\rangle\rangle$ n'est absolument pas indispensable. En effet, pour établir la formule $\delta z_n(x, y) = z_n(x', y') + z_n(x'', y'')$, on peut reproduire les calculs (3) en considérant non pas des séries formelles, mais certaines de leurs portions assez longues (en fonction de n) et en se concentrant uniquement sur les termes de degré $\leqslant n$. Il est clair que le calcul (qui sera réalisé désormais à l'intérieur de l'algèbre $K\langle x, y\rangle$) reste entièrement en vigueur. On peut donc considérer en définitive que l'assertion B est prouvée.

La démonstration de l'assertion B qui a été produite ne nous donne aucune formule explicite pour les polynômes de Lie $\mathfrak{z}_n(x, y)$ dont elle affirme l'existence. On se propose de remédier à cette situation.

Supposons comme dans la leçon précédente que $K^*\langle x, y\rangle$ est la sous-algèbre de l'algèbre $K\langle x, y\rangle$, formée de tous les polynômes

sans terme constant. Définissons une application

$$\sigma : K^*\langle x, y\rangle \to \mathfrak{l}\langle x, y\rangle$$

satisfaisant aux conditions suivantes:

a) l'application σ est linéaire;

b) $\sigma x = x$ et $\sigma y = y$;

c) pour tout monôme $a \neq 1$

$$\sigma(ax) = [\sigma a, x], \quad \sigma(ay) = [\sigma a, y].$$

Il est évident que ces conditions définissent une seule application σ.

Définition 1. Les éléments de l'algèbre $\mathfrak{l}\langle x, y\rangle$ de la forme σa, où a est un monôme, s'appellent *monômes de Lie*.

Par définition:

1) les éléments x et y sont des monômes de Lie;

2) si u est un monôme de Lie, il en est de même de $[u, x]$ et $[u, y]$.

Il est évident que l'enveloppe linéaire de tous les monômes de Lie est confondue avec $\operatorname{Im}\sigma = \sigma(K^*\langle x, y\rangle)$.

L'expression explicite du monôme de Lie σa associé au monôme $a = x^{p_1}y^{q_1} \ldots x^{p_k}y^{q_k}$ est

$$(4) \qquad \sigma a = [\ldots\underbrace{[x, x], x], \ldots, x]}_{p_1 \text{ fois}}, \underbrace{y], \ldots, y]}_{q_1 \text{ fois}}, \ldots$$

$$\ldots, \underbrace{x], \ldots, x]}_{p_k \text{ fois}}, \underbrace{y], \ldots, y]}_{q_k \text{ fois}}.$$

On voit, en particulier, que $\sigma a = 0$ si $p_1 > 1$ ou si $p_1 = 0$ et $q_1 > 1$ (on rappelle que d'après la définition du monôme, tous les exposants $p_1, q_1, \ldots, p_k, q_k$ sont strictement positifs, hormis, éventuellement, p_1 et q_k).

La formule (4) entraîne immédiatement que *pour tout monôme de Lie $u = \sigma a$, le polynôme ιu est un élément homogène de l'algèbre* $K\langle x, y\rangle$, c'est-à-dire une combinaison linéaire de monômes de même degré $n \geqslant 1$. Ce degré (qui est visiblement égal à celui du monôme a) sera appelé *degré* du monôme de Lie u.

Lemme 1. *Pour tous monômes de Lie u et v, l'élément* $[u, v]$ *appartient à l'espace vectoriel* $\operatorname{Im}\sigma$.

D é m o n s t r a t i o n. Raisonnons par récurrence sur le degré n de v. Si $n = 1$, c'est-à-dire si $v = x$ ou $v = y$, l'élément $[u, v]$ est par définition un monôme de Lie, donc est contenu dans $\operatorname{Im}\sigma$. Supposons que $n > 1$. Alors $v = [w, x]$ ou $v = [w, y]$, où w est un monôme de Lie de degré $n - 1$. Supposons pour fixer les idées que $v = [w, x]$ (le cas $v = [w, y]$ est exactement le même). Alors $[u, v] = [u, [w, x]]$ et, par suite, en vertu de l'identité de Jacobi,

$$[u, v] = [[u, w], x] - [[u, x], w].$$

Or, d'après l'hypothèse de récurrence, l'élément $[u, w]$, donc $[[u, w], x]$ appartient à $\operatorname{Im} \sigma$, c'est-à-dire est une combinaison linéaire de monômes de Lie. Il en est de même de l'élément $[[u, x], w]$. Donc $[u, v] \in \operatorname{Im} \sigma$. □

Corollaire. *Tout élément de l'algèbre de Lie $\mathfrak{l}\langle x, y\rangle$ est une combinaison linéaire de monômes de Lie. En d'autres termes,*

$$\operatorname{Im} \sigma = \mathfrak{l}\langle x, y\rangle.$$

D é m o n s t r a t i o n. Du lemme 1, il résulte immédiatement que l'espace $\operatorname{Im} \sigma$ est une sous-algèbre de l'algèbre de Lie $\mathfrak{l}\langle x, y\rangle$. Comme $x, y \in \operatorname{Im} \sigma$ et que l'algèbre $\mathfrak{l}\langle x, y\rangle$ est engendrée par les éléments x et y, cet espace doit être confondu avec $\mathfrak{l}\langle x, y\rangle$. □

Définissons maintenant un *antihomomorphisme* de l'algèbre $\mathfrak{K}\langle x, y\rangle$ dans l'algèbre $\operatorname{End}_{\text{lin}}(\mathfrak{l}\langle x, y\rangle)$ des applications linéaires de l'algèbre $\mathfrak{l}\langle x, y\rangle$ dans elle-même, c'est-à-dire une application linéaire

$$\theta : \mathfrak{K}\langle x, y\rangle \to \operatorname{End}_{\text{lin}}(\mathfrak{l}\langle x, y\rangle),$$

satisfaisant à la relation $\theta(ab) = \theta(b) \circ \theta(a)$ quels que soient $a, b \in \mathfrak{K}\langle x, y\rangle$. L'algèbre $\mathfrak{K}\langle x, y\rangle$ étant engendrée par les éléments x et y, cet antihomomorphisme est défini de façon unique par ses valeurs θx et θy, et comme les générateurs x et y sont des générateurs libres, les applications θx et θy peuvent être arbitrairement choisies. Définissons-les à l'aide des formules

$$(\theta x)\, v = [v, x], \quad (\theta y)\, v = [v, y],$$

où v est un élément arbitraire de l'algèbre $\mathfrak{l}\langle x, y\rangle$.

Donc, d'après cette définition, si $u = x$ ou $u = y$, l'application $\theta(u) : \mathfrak{l}\langle x, y\rangle \to \mathfrak{l}\langle x, y\rangle$ est une dérivation intérieure $\operatorname{ad}(-u) = -\operatorname{ad} u$ de l'algèbre de Lie $\mathfrak{l}\langle x, y\rangle$ (cf. leçon 3). Il s'avère que l'égalité $\theta u = -\operatorname{ad} u$ (ou, plus exactement, l'égalité $\theta(\iota u) = -\operatorname{ad} u$) est valable pour tout élément u de l'algèbre $\mathfrak{l}\langle x, y\rangle$.

Lemme 2. *Pour tous éléments $u, v \in \mathfrak{l}\langle x, y\rangle$, on a*

$$(\bar{\theta} u)\, v = [v, u], \tag{5}$$

où $\bar{\theta} = \theta \circ \iota$.

D é m o n s t r a t i o n. D'après le corollaire du lemme 1, il suffit de prouver l'égalité (5) uniquement pour le cas où l'élément u est un monôme de Lie. Raisonnons par récurrence sur le degré n de l'élément u. Si $n = 1$, alors $u = x$ ou $u = y$, et l'égalité (5) est réalisée par définition. Soit $n > 1$. Alors $u = [w, x]$ ou $u = [w, y]$, où w est un monôme de Lie de degré $n - 1$. Supposons pour fixer les idées que $u = [w, x]$. Comme

$$\bar{\theta}(w, x) = \theta(bx) - \theta(xb) = \theta x \cdot \theta b - \theta b \cdot \theta x,$$

où $b = \iota w$, on a en vertu de l'hypothèse de récurrence et de l'identité de Jacobi

$$(\bar{\theta} u)\, v = (\bar{\theta}\, [w,\ x])\, v = (\theta x)\, (\theta b)\, v - (\theta b)\, (\theta x)\, v =$$
$$= (\bar{\theta} x)\, (\bar{\theta} w)\, v - (\bar{\theta} w)\, (\bar{\theta} x)\, v =$$
$$= (\bar{\theta} x)\, [v,\ w] - (\bar{\theta} w)\, [v,\ x] = [[v,\ w],\ x] - [[v,\ x],\ w] =$$
$$= [v,\ [w,\ x]] = [v,\ u]. \quad \square$$

Lemme 3. *Pour tous éléments a, $b \in K^*\langle x, y\rangle$, on a*

$$\sigma\,(ab) = (\theta b)\,(\sigma a). \tag{6}$$

D é m o n s t r a t i o n. On peut sans perdre en généralité admettre que b est un monôme. Raisonnons par récurrence sur son degré n. Pour $n = 1$, c'est-à-dire si $b = x$ ou $b = y$, l'égalité (6) résulte immédiatement des définitions. Soit $n > 1$. Alors $b = cx$ ou $b = cy$, où c est un monôme de degré $n - 1$. Supposons pour fixer les idées que $b = cx$. Comme

$$\sigma\,(ab) = \sigma\,(acx) = [\sigma\,(ac),\ x] = (\theta x)\,(\sigma\,(ac)),$$

il vient d'après l'hypothèse de récurrence

$$\sigma\,(ab) = (\theta x)\,(\theta c)\,(\sigma a) = \theta\,(cx)\,(\sigma a) = (\theta b)\,(\sigma a). \quad \square$$

Lemme 4. *La restriction $\bar{\sigma} = \sigma \circ \iota$ de l'application σ à $l\langle x, y\rangle \subset K^*\langle x, y\rangle$ est une dérivation de l'algèbre $l\langle x, y\rangle$, c'est-à-dire que pour tous éléments u, $v \in l\langle x, y\rangle$, on a*

$$\bar{\sigma}\,[u,\ v] = [\bar{\sigma} u,\ v] + [u,\ \bar{\sigma} v].$$

D é m o n s t r a t i o n. Soient $\iota u = a$ et $\iota v = b$. Il vient en vertu des lemmes 3 et 2

$$\bar{\sigma}\,[u,\ v] = \sigma\,(ab) - \sigma\,(ba) = (\theta b)\,(\sigma a) - (\theta a)\,(\sigma b) =$$
$$= [\sigma a,\ b] - [\sigma b,\ a] = [\sigma a,\ b] + [a,\ \sigma b] =$$
$$= [\bar{\sigma} u,\ v] + [u,\ \bar{\sigma} v]. \quad \square$$

Nous pouvons prouver maintenant la proposition fondamentale.

Proposition 2 (théorème de Dynkine). *Un polynôme homogène $a \in K\langle x, y\rangle$ de degré $n \geqslant 1$ est un élément de l'algèbre de Lie $l\langle x, y\rangle$ (c'est-à-dire est de la forme ιu, où $u \in l\langle x, y\rangle$) si et seulement si*

$$\iota\,(\sigma a) = na. \tag{7}$$

D é m o n s t r a t i o n. Si l'égalité (7) est réalisée, alors

$$a = \iota\left(\frac{\sigma a}{n}\right), \quad \text{où} \quad \frac{\sigma a}{n} \in l\,\langle x,\ y\rangle.$$

Réciproquement, soit $a = \iota u$ où $u \in \mathfrak{l}\langle x, y\rangle$. Sans nuire à la généralité, on peut admettre que l'élément u est un monôme de Lie. Effectuons une récurrence sur le degré n de l'élément u. Si $n = 1$, alors $u = x$ ou $u = y$ et l'égalité (7) est remplie. Soit $n > 1$ et supposons, pour fixer les idées, que $u = [v, x]$, où v est un monôme de Lie de degré $n - 1$. D'après le lemme 4 et l'hypothèse de récurrence, on a alors

$$\begin{aligned}\iota(\sigma a) = \iota(\bar{\sigma} u) &= \iota\bar{\sigma}[v, x] = \\ &= \iota([\bar{\sigma} v, x] + [v, \bar{\sigma} x]) = \\ &= [\iota\bar{\sigma} v, x] + \iota[v, x] = \\ &= (n-1)\iota[v, x] + \iota[v, x] = \\ &= (n-1)\iota u + \iota u = (n-1)a + a = na. \ \square\end{aligned}$$

La proposition 2 dit que si $a = \iota u$, alors

$$u = \frac{\sigma a}{n}.$$

D'après l'assertion B, la condition $a = \iota u$ est réalisée pour $a = z_n(x, y)$ et $u = \mathfrak{Z}_n(x, y)$. Donc

$$\mathfrak{Z}_n(x, y) = \frac{\bar{\sigma} \bar{z}_n(x, y)}{n},$$

c'est-à-dire que (cf. formule (2))

$$(8) \qquad \mathfrak{Z}_n(x, y) = \frac{1}{n}\sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \sum_{(p),(q)} \frac{\sigma\left(x^{p_1}y^{q_1}\dots x^{p_k}y^{q_k}\right)}{p_1!\,q_1!\dots p_k!\,q_k!},$$

où dans la somme intérieure, la sommation est étendue à tous les exposants entiers strictement positifs $p_1, q_1, \dots, p_k, q_k$ tels que

$$p_1 + q_1 > 0, \dots, p_k + q_k > 0$$

et

$$p_1 + q_1 + \dots + p_k + q_k = n.$$

Ceci est la formule désirée.

Pour la série formelle de Lie

$$(9) \qquad \mathfrak{Z}(x, y) = \sum_{n=1}^{\infty} \mathfrak{Z}_n(x, y)$$

on obtient, en regroupant les termes, la formule suivante

$$(10) \qquad \mathfrak{Z}(x, y) = \sum_{k=1}^{\infty} \sum_{(p),(q)} \frac{(-1)^{k-1}}{k} \cdot \frac{1}{p_1+q_1+\dots+p_k+q_k} \times \\ \times \frac{[x^{p_1}y^{q_1}\dots x^{p_k}y^{q_k}]}{p_1!\,q_1!\dots p_k!\,q_k!},$$

où pour la suggestion on a substitué $[x^{p_1}y^{q_1} \ldots x^{p_k}y^{q_k}]$ à $\sigma(x^{p_1}y^{q_1} \ldots \ldots x^{p_k}y^{q_k})$.

Cette série s'appelle *série de Campbell-Hausdorff sous la forme de Dynkine.*

La série (10) est peu commode pour les calculs pratiques, car elle contient de nombreux termes semblables non réduits. La difficulté à réduire ces termes est bien illustrée par le calcul de la partie linéaire en x de la série (10).

Un monôme de Lie $[x^{p_1}y^{q_1} \ldots x^{p_m}y^{q_m}]$ est linéaire en x (et non nul) si et seulement si ou bien $p_1 = 1, p_2 = 0, \ldots, p_m = 0$ (et alors il est de la forme $[xy^n]$, où $n = q_1 + \ldots + q_m$), ou bien $p_1 = 0$, $p_2 = 1$, $p_3 = 0, \ldots, p_m = 0$, $q_1 = 1$ (et alors il est de la forme $[yxy^{n-1}]$, où $n - 1 = q_2 + \ldots + q_m$). Sous ces conditions, dans le premier cas $q_1 \geqslant 0$, $q_2 \geqslant 1, \ldots, q_m \geqslant 1$, et dans le second, $q_2 \geqslant 0$, $q_3 \geqslant 1, \ldots, q_m \geqslant 1$. Comme $[yxy^{n-1}] = -[xy^n]$, on voit — en changeant les notations — que la partie linéaire en x de la série $\mathfrak{Z}(x, y)$ (qui sera désignée par $\mathfrak{Z}'(x, y)$) est exprimée par la formule

$$\mathfrak{Z}'(x, y) = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} \sum_{n=m-1}^{\infty} \frac{1}{n+1} \times$$

$$\times \left(\sum{}' \frac{1}{q_1!\, q_2! \ldots q_m!} - \sum{}'' \frac{1}{q_1!\, q_2! \ldots q_{m-1}!} \right) [xy]^n,$$

où dans $\sum'$ la sommation est étendue à toutes les collections d'entiers $q_1, q_2, \ldots, q_m$ tels que

$$q_1 \geqslant 0,\ q_2 \geqslant 1, \ldots, q_m \geqslant 1 \quad \text{et} \quad q_1 + q_2 + \ldots + q_m = n,$$

et dans $\sum''$, à $q_1, q_2, \ldots, q_{m-1}$ tels que

$$q_1 \geqslant 0,\ q_2 \geqslant 1, \ldots, q_{m-1} \geqslant 1$$

et

$$q_1 + q_2 + \ldots + q_{m-1} = n - 1.$$

La somme $\sum''$ disparaît pour $m = 1$.

Pour calculer les coefficients de la série $\mathfrak{Z}'(x, y)$, on se servira de la méthode standard des fonctions génératrices. Soit $A(t)$ la série déduite de $\mathfrak{Z}'(x, y)$ en substituant t^n à $[xy^n]$. Comme

$$\sum_{n=m-1}^{\infty} \sum{}' \frac{1}{q_1!\, q_2! \ldots q_m!} \frac{t^n}{n+1} = \frac{1}{t} \sum_{n=m-1}^{\infty} \int_0^t dt \sum{}' \frac{1}{q_1!\, q_2! \ldots q_m!} t^n =$$

$$= \frac{1}{t} \int_0^t dt \left(\sum_{q_1=0}^{\infty} \frac{t^{q_1}}{q_1!} \right) \left(\sum_{q_2=1}^{\infty} \frac{t^{q_2}}{q_2!} \right)^{m-1} = \frac{1}{t} \int_0^t e^t (e^t - 1)^{m-1}\, dt$$

et, de façon analogue,

$$\sum_{n=m-1}^{\infty} \sum{}'' \frac{1}{q_1!\, q_2! \dots q_{m-1}!} \frac{t^n}{n+1} = \frac{1}{t} \int_0^t te^t (e^t - 1)^{m-2}\, dt,$$

il vient

$$A(t) = \frac{1}{t} \int_0^t \frac{e^t}{e^t - 1} \left(\sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} (e^t - 1)^m \right) dt -$$

$$- \frac{1}{t} \int_0^t \frac{te^t}{(e^t - 1)^2} \left(\sum_{m=2}^{\infty} \frac{(-1)^{m-1}}{m} (e^t - 1)^m \right) dt =$$

$$= \frac{1}{t} \int_0^t \frac{e^t}{e^t - 1}\, t\, dt - \frac{1}{t} \int_0^t \frac{te^t}{(e^t - 1)^2} (t - (e^t - 1))\, dt =$$

$$= \frac{te^t}{e^t - 1} = \frac{-t}{e^{-t} - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} (-t)^n,$$

où B_n sont les nombres de Bernoulli.

Comme $[xy^n] = (-\mathrm{ad}\ y)^n x$, ceci prouve que

$$\mathfrak{H}'(x, y) = \sum_{n=0}^{\infty} \frac{B_n}{n!} (\mathrm{ad}\ y)^n x = \frac{\mathrm{ad}\ y}{e^{\mathrm{ad}\ y} - 1} x,$$

c'est-à-dire (puisque y est le seul terme de la série $\mathfrak{H}(x, y)$ à ne pas contenir x) que

(11) $$\mathfrak{H}(x, y) = y + \frac{\mathrm{ad}\ y}{e^{\mathrm{ad}\ y} - 1} x + \dots,$$

où les points de suspension figurent les termes dont le degré de x est $\geqslant 2$.

On montre de façon analogue que

(12) $$\mathfrak{H}(x, y) = x + \frac{-\mathrm{ad}\ x}{e^{-\mathrm{ad}\ x} - 1} y + \dots,$$

où les points de suspension figurent les termes dont le degré de y est $\geqslant 2$.

Etudions maintenant la convergence des séries (9) et (10). Utilisons à cet effet le lemme 1 de la leçon 2, qui, comme déjà signalé, s'applique à toute algèbre de dimension finie et, en particulier, à toute algèbre de Lie $\mathfrak{g}$ de dimension finie.

Soient $\|\ \|$ une norme multiplicative sur $\mathfrak{g}$, X et Y des éléments de $\mathfrak{g}$ tels que $\|X\| < \delta$ et $\|Y\| < \delta$, où $0 < \delta < 1$.

L'algèbre $\mathfrak{l}\langle x, y\rangle$ étant libre, il existe un seul homomorphisme $\mathfrak{l}\langle x, y\rangle \to \mathfrak{g}$ transformant x et y en X et Y. L'image d'un élément $u = u(x, y)$ de l'algèbre $\mathfrak{l}\langle x, y\rangle$ par cet homomorphisme sera notée $u(X, Y)$.

Une récurrence immédiate (qui utilise la multiplicativité de la norme) montre maintenant que si u est un monôme de Lie de degré n, alors

$$\| u(X, Y) \| \leqslant \delta^n.$$

De là il s'ensuit que si $u = \sum c_\alpha u_\alpha$, où u_α sont des monômes de Lie de degré n, alors

$$\| u(X, Y) \| \leqslant C\delta^n,$$

où $C = \sum_\alpha | c_\alpha |$. On obtient, en particulier (cf. (8)),

$$(13) \qquad \| \mathfrak{z}_n(X, Y) \| \leqslant D_n\delta^n,$$

où

$$D_n = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \sum_{(p),(q)} \frac{1}{p_1!\, q_1! \ldots p_k!\, q_k!}.$$

La somme intérieure prend une valeur égale au coefficient de t^n dans la série de Maclaurin de la fonction $(e^{2t} - 1)^k$. Donc, le nombre nD_n est égal au coefficient de t^n dans la série de Maclaurin de la fonction $\sum_{k=1}^{n} \frac{1}{k} (e^{2t} - 1)^k$ ou, ce qui est équivalent, de la fonction

$$f(t) = \sum_{k=1}^{\infty} \frac{1}{k} (e^{2t} - 1)^k.$$

Donc,

$$(14) \qquad \sum_{n=1}^{\infty} D_n\delta^n = \int_0^\delta \frac{f(t)}{t}\, dt.$$

Comme la série de $f(t)$ converge manifestement pour $| e^{2t} - 1 | < 1$ et donc pour $| t | < \frac{\ln 2}{2}$, ceci prouve que *la série* (14) *converge pour* $\delta < \frac{\ln 2}{2}$. Puisque, en vertu de la formule (13), la série (14) majore la série

$$(15) \qquad \mathfrak{z}(X, Y) = \sum_{n=1}^{\infty} \mathfrak{z}_n(X, Y)$$

il vient que *la série* (15) *converge si* $\| X \| < \delta_0$, $\| Y \| < \delta_0$, *où* $\delta_0 = \frac{\ln 2}{2}$. □

Remarque 1. La convergence de la série (15) pour $\| X \| < \delta_0$ et $\| Y \| < \delta_0$ a déjà été prouvée dans la leçon 4. Le fait nouveau, c'est que la convergence de la série (15) a été acquise désormais pour toute algèbre de Lie $\mathfrak{g}$ (de dimension finie), tandis que dans la leçon 5 l'algèbre $\mathfrak{g}$ était l'algèbre de Lie d'un groupe (groupuscule) de Lie.

Soient maintenant G un groupe (ou un groupuscule) de Lie analytique, $\mathfrak{g} = \mathfrak{l}(G)$, son algèbre de Lie. L'application exponentielle exp étant un difféomorphisme au point $0 \in \mathfrak{g}$, elle permet de transporter (au moyen de la formule $X \cdot Y = \exp^{-1}(\exp X \cdot \exp Y)$) la multiplication de G sur un voisinage de $0 \in \mathfrak{g}$. Donc, l'espace vectoriel $\mathfrak{g}$ se transforme en un groupuscule de Lie, isomorphe (dans la catégorie GR-LOC) au groupuscule G (l'isomorphisme est réalisé par l'application exp).

Donc, outre les opérations linéaires et le crochet de Lie $X, Y \mapsto [X, Y]$, l'espace vectoriel $\mathfrak{g}$ est muni d'une opération (la « multiplication ») pour laquelle $\mathfrak{g}$ est un groupuscule de Lie. Cette opération est liée aux opérations de l'algèbre de Lie $\mathfrak{g}$ par la formule

$$X \cdot Y = \mathcal{H}(X, Y).$$

L'objet construit mérite d'être spécialement défini.

Définition 2. On appellera *groupalgèbre de Lie* un espace vectoriel $\mathfrak{g}$ de dimension finie sur le corps $\mathbb{R}$, muni d'une structure d'algèbre de Lie (avec la multiplication $[x, y]$) et d'une structure de groupuscule de Lie (avec la multiplication xy), et dans lequel le produit xy est défini si et seulement si la série de Campbell-Hausdorff $\mathcal{H}(x, y)$ converge, auquel cas

$$xy = \mathcal{H}(x, y).$$

Des formules (11) et (12) il s'ensuit immédiatement que pour tout élément X (assez proche de 0) d'un groupalgèbre de Lie $\mathfrak{g}$, les différentielles $(dL_X)_0$ et $(dR_Y)_0$ des translations $L_X : Y \mapsto \mathcal{H}(X, Y)$ et $R_Y : X \mapsto \mathcal{H}(X, Y)$ au point 0 sont définies par les formules

$$(dL_X)_0 = \frac{-\operatorname{ad} X}{e^{-\operatorname{ad} X} - E}, \quad (dR_X)_0 = \frac{e^{\operatorname{ad} X}}{e^{\operatorname{ad} X} - E}.$$

Remarque 2. La deuxième formule a déjà été prouvée par une autre méthode à la leçon 4 (cf. corollaire 1 de la proposition 3 de la leçon 4).

Une application $\mathfrak{g} \to \mathfrak{h}$ de groupalgèbres de Lie est par définition un *homomorphisme* si elle est un homomorphisme d'algèbres de Lie (donc de groupuscules de Lie).

Il est évident que les groupalgèbres de Lie et leurs homomorphismes forment une catégorie. Nous noterons cette catégorie GRIE.

D'après ce qui précède, à tout groupuscule de Lie G on peut associer un groupalgèbre de Lie en transportant la multiplication de G sur $\mathfrak{g}$ au moyen de l'application exponentielle exp. Nous désignerons ce groupalgèbre de Lie par $\mathfrak{l}'(G)$. Il est évident que la correspondance $G \mapsto \mathfrak{l}'(G)$ est un foncteur

$$\mathfrak{l}'\colon \text{GR-LOC} \to \text{GRIE}.$$

Ce foncteur sera aussi appelé *foncteur de Lie.*

Considérons maintenant le *foncteur d'annihilation*

$$I\colon \text{GRIE} \to \text{GR-LOC},$$

dont l'action consiste à annihiler la structure d'algèbre de Lie de chaque groupalgèbre de Lie.

De prime abord, les foncteurs $\mathfrak{l}'$ et I semblent être réciproques l'un de l'autre. Mais ce n'est qu'une impression. En effet, pour tout groupuscule de Lie G, le groupuscule $(I \circ \mathfrak{l}')\, G$ n'est pas le groupuscule G, mais le groupuscule construit ci-dessus sur l'espace vectoriel $\mathfrak{l}(G)$ qui est généralement différent de G. Cependant on voit que le groupuscule $(I \circ \mathfrak{l}')\, G$ est canoniquement isomorphe au groupuscule G (l'isomorphisme correspondant est l'application exponentielle exp). Dans le langage de la théorie des foncteurs, cela signifie que *le foncteur* $I \circ \mathfrak{l}'$ *est isomorphe au foncteur identique* Id *de la catégorie* GR-LOC:

$$I \circ \mathfrak{l} \approx \text{Id}.$$

De façon analogue, *le foncteur* $\mathfrak{l}' \circ I$ *est isomorphe au foncteur identique* Id *de la catégorie* GRIE:

$$\mathfrak{l}' \circ I \approx \text{Id}.$$

En effet, le foncteur $\mathfrak{l}' \circ I$ associe au groupalgèbre de Lie $\mathfrak{g}$ un groupalgèbre qui est confondu, en tant qu'espace vectoriel, avec l'espace tangent $\mathsf{T}_0(\mathfrak{g})$. Or, nous savons qu'entre les espaces vectoriels $\mathfrak{g}$ et $\mathsf{T}_0(\mathfrak{g})$, il existe un isomorphisme canonique

$$l\colon \mathfrak{g} \to \mathsf{T}_0(\mathfrak{g}).$$

Montrons que l est un isomorphisme de groupalgèbres de Lie.

Par définition, l'isomorphisme l associe à tout élément $x \in \mathfrak{g}$ le vecteur tangent à la courbe $t \mapsto tx$ au point 0. Comme $(tx)\cdot(sx) = \mathfrak{H}(tx, sx) = \mathfrak{H}_1(tx, sx) = (t+s)\,x$, cette courbe est un sous-groupe à un paramètre du groupuscule de Lie $I\mathfrak{g}$, donc (cf. proposition 1 de la leçon 4) pour tous éléments $x, y \in \mathfrak{g}$, le vecteur $[lx, ly]$ est un vecteur tangent en 0 à la courbe

$$t \mapsto (\sqrt{t}\,x)\,(\sqrt{t}\,y)\,(\sqrt{t}\,x)^{-1}\,(\sqrt{t}\,y)^{-1}.$$

Comme (cf. démonstration de la proposition 1 de la leçon 4)

$$(\sqrt{t}x)(\sqrt{t}y)(\sqrt{t}x)^{-1}(\sqrt{t}y)^{-1} = t\,[x, y] + O\,(t^{3/2}),$$

cette courbe admet en 0 le même vecteur tangent que la courbe $t \mapsto t\,[x, y]$. Donc,

$$[lx, ly] = l\,[x, y],$$

de sorte que l est bien un isomorphisme d'algèbres et, par suite, de groupalgèbres de Lie. □

Si des foncteurs $F\colon \mathsf{C} \to \mathsf{D}$ et $G\colon \mathsf{D} \to \mathsf{C}$, où C et D sont des catégories, sont tels que les composés $F \circ G$ et $G \circ F$ sont isomorphes aux foncteurs identiques (des catégories D et C respectivement) on dit qu'ils sont *quasiréciproques.* Les catégories C et D pour lesquelles existent des foncteurs quasiréciproques $\mathsf{C} \to \mathsf{D}$ et $\mathsf{D} \to \mathsf{C}$ sont dites *équivalentes.*

On a donc prouvé la proposition suivante:

Proposition 3. *Les catégories* GR-LOC *et* GRIE *sont équivalentes, l'équivalence étant réalisée par les foncteurs quasiréciproques* $\mathfrak{l}'$ *et* I. □

En ignorant la structure de groupuscule du groupalgèbre de Lie, on obtient le foncteur d'annihilation

$$J\colon \text{GRIE} \to \text{ALG}_f\text{-LIE};$$

le composé $J \circ \mathfrak{l}'$ de ce foncteur et du foncteur de Lie $\mathfrak{l}'\colon$ GR-LOC → GRIE n'est autre que le foncteur de Lie de groupuscules

$$\mathfrak{l}\colon \text{GR-LOC} \to \text{ALG}_f\text{-LIE}$$

qui nous intéresse en premier lieu.

Proposition 4. *Le foncteur J admet le foncteur réciproque*

$$J^{-1}\colon \text{ALG}_f\text{-LIE} \to \text{GRIE}.$$

D é m o n s t r a t i o n. Soit $\mathfrak{g}$ une algèbre de Lie (dont la multiplication est notée $[x, y]$). Si le groupalgèbre $J^{-1}\mathfrak{g}$ existe, il est confondu en tant qu'algèbre de Lie avec l'algèbre $\mathfrak{g}$; quant à sa multiplication, elle est liée avec les opérations dans $\mathfrak{g}$ par la formule

$$xy = \mathfrak{H}(x, y).$$

Donc, pour prouver la proposition 3, il suffit de montrer que pour toute algèbre de Lie $\mathfrak{g}$, cette formule définit dans $\mathfrak{g}$ une multiplication satisfaisant aux axiomes d'un groupuscule de Lie. Pour cela il suffit de démontrer que:

a) dans l'algèbre $\mathfrak{g}$ il existe un voisinage U de 0 dans lequel converge la série $\mathfrak{H}(x, y)$, $x, y \in U$;

b) l'application $U \times U \to \mathfrak{g}$, induite par la correspondance $x, y \mapsto \mathfrak{H}(x, y)$, et l'application $x \mapsto x^{-1} = -x$, possèdent les propriétés énumérées dans la définition 1 de la leçon 4.

Nous avons établi plus haut l'affirmation a); pour prouver l'affirmation b), on remarquera tout d'abord que dans l'algèbre des séries formelles de trois variables x, y et z non permutables, on a la formule

$$(e^x e^y)\, e^z = e^x\, (e^y e^z)$$

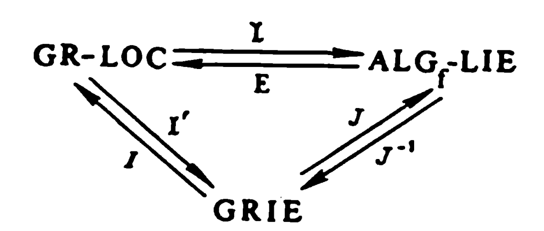
(dont la vérification est immédiate). Comme la série de Campbell-Hausdorff $\mathfrak{H}(x, y)$ se transforme par la substitution $[x, y] \mapsto xy - yx$ en la série $\ln(e^x e^y)$, il s'ensuit immédiatement de là que la multiplication $x, y \mapsto \mathfrak{H}(x, y)$ est associative (si seulement les séries concernées convergent). De façon analogue, puisque $\mathfrak{H}(0, x) = \mathfrak{H}(x, 0) = 0$, la multiplication $x, y \mapsto \mathfrak{H}(x, y)$ admet 0 pour élément neutre, et comme $e^x e^{-x} = 1$, l'élément réciproque x^{-1} relativement à cette multiplication est l'élément $-x$.

Ceci achève la démonstration de la proposition 4. □

Le foncteur I étant quasiréciproque du foncteur $\mathfrak{l}'$, le foncteur $E = I \circ J^{-1}$ est quasiréciproque du foncteur de Lie $\mathfrak{l} = J \circ \mathfrak{l}'$. Nous avons donc atteint notre objectif principal: nous avons trouvé un foncteur qui, à défaut d'être réciproque, est quasiréciproque du foncteur de Lie GR-LOC → ALG_f-LIE.

Les résultats acquis peuvent être résumés par le théorème suivant.

Théorème 1. *On a le diagramme commutatif suivant dans lequel*



les foncteurs J et J^{-1} sont réciproques l'un de l'autre, et les foncteurs $\mathfrak{l}$ et E, d'une part, et les foncteurs $\mathfrak{l}'$ et I, de l'autre, sont quasiréciproques. □

Corollaire. *Les catégories* GR-LOC, ALG_f-LIE *et* GRIE *sont équivalentes.* □

Dans le fond le théorème 1 était déjà connu de Sophus Lie.

Lie a exposé ses résultats sous forme de six théorèmes: trois directs et trois réciproques. Le plus proche du théorème 1 était son troisième théorème réciproque. Pour cette raison boiteuse, le théorème 1 est parfois appelé *troisième théorème de Lie*.

A noter que dans le théorème 1, les groupuscules de Lie sont supposés être a n a l y t i q u e s. Le cas de groupuscules appartenant à la classe C^r, $2 \leqslant r \leqslant \infty$, sera traité dans la leçon prochaine.

LEÇON 7

Sous-groupuscules et sous-algèbres.— Sous-groupuscules invariants et idéaux.— Groupuscules quotients et algèbres quotients.— Réduction de groupuscules différentiables à des groupuscules analytiques.— Systèmes de Pfaff.— Sous-fibrés de fibrés tangents.— Sous-fibrés intégrables.— Graphes de systèmes de Pfaff.— Sous-fibrés involutifs.—Univalence complète d'un foncteur de Lie.— Involutivité des sous-fibrés intégrables.— Sous-fibrés complètement intégrables.

Le théorème 1 de la leçon précédente ramène intégralement la théorie des groupuscules de Lie (analytiques) à la théorie des algèbres de Lie. Nous allons illustrer ceci sur l'exemple des sous-groupuscules et des groupuscules quotients.

Définition 1. On appelle *sous-groupuscule* d'un groupuscule de Lie G un sous-ensemble H de G tel que:

1) si pour des éléments $a, b \in H$ est défini l'élément ab^{-1}, alors $ab^{-1} \in H$;

2) il existe un voisinage U de l'unité e du groupuscule G tel que l'intersection $U \cap H$ est fermée dans U (*condition de fermeture locale*).

A noter que de cette définition il ne s'ensuit pas encore avec évidence que le sous-groupuscule H est lui-même un groupuscule de Lie (bien que comme nous le verrons plus bas ce soit le cas).

Soient $\mathfrak{g}$ une algèbre de Lie, $\mathfrak{h}$ une sous-algèbre de $\mathfrak{g}$. En appliquant le foncteur $E = I \circ J^{-1}$ à $\mathfrak{g}$ et à $\mathfrak{h}$, on peut construire les groupuscules de Lie $E\mathfrak{g}$ et $E\mathfrak{h}$. Par définition, $E\mathfrak{h}$ sera contenu dans $E\mathfrak{g}$ et, de plus, sera son sous-groupuscule (puisque si $x, y \in \mathfrak{h}$, alors $\mathfrak{Z}_n(x, y) \in \mathfrak{h}$ pour tout $n \geqslant 1$).

Soient maintenant G un groupuscule de Lie, $\mathfrak{g} = \mathfrak{l}(G)$ son algèbre de Lie. L'isomorphisme canonique $E\mathfrak{g} \approx G$ associe à tout sous-groupuscule du groupuscule $E\mathfrak{g}$, de la forme $E\mathfrak{h}$, où $\mathfrak{h}$ est une sous-algèbre de $\mathfrak{g}$, un sous-groupuscule de G. Ce sous-groupuscule sera confondu dans un voisinage de l'unité $e \in G$ avec l'image $\exp \mathfrak{h}$ de la sous-algèbre $\mathfrak{h}$ par l'application exponentielle exp, de sorte

que nous pouvons le noter exp $\mathfrak{h}$, conformément à notre convention de ne pas distinguer des groupuscules équivalents.

Le sous-groupuscule exp $\mathfrak{h}$ de G associé à la sous-algèbre $\mathfrak{h}$ de $\mathfrak{g} = \mathfrak{l}(G)$ est tel que dans un système de coordonnées normales $x^1, \ldots, x^n$ (plus exactement, dans un système défini par une base de $\mathfrak{g}$ dont les m premiers vecteurs forment une base de $\mathfrak{h}$) il est défini par les équations linéaires

$$x^{m+1} = 0, \ldots, x^n = 0. \tag{1}$$

Les sous-groupuscules jouissant de cette propriété seront appelés *localement plans*. Donc, pour qu'un sous-groupuscule $H \subset G$ soit de la forme exp $\mathfrak{h}$, il est nécessaire qu'il soit localement plan. Il est aisé de voir que cette condition nécessaire est aussi suffisante, de sorte qu'*un sous-groupuscule H d'un groupuscule de Lie G est associé à une sous-algèbre de l'algèbre de Lie* $\mathfrak{g} = \mathfrak{l}(G)$ (donc, en particulier, est un groupuscule de Lie) *si et seulement s'il est localement plan*. En effet, soit $\mathfrak{h}$ le sous-espace de $\mathfrak{g}$ engendré par les m premiers vecteurs de la base définissant les coordonnées normales dans lesquelles H est défini par les équations linéaires (1). Alors pour tout $X \in \mathfrak{h}$ et tout t (tel que $|t|$ soit assez petit), le point $\beta_X(t) = \exp tX$ appartient à H. Donc, si $X, Y \in \mathfrak{h}$, alors

$$\beta(t) = \beta_X(\sqrt{t})\,\beta_Y(\sqrt{t})\,\beta_X(\sqrt{t})^{-1}\beta_Y(\sqrt{t})^{-1} \in H,$$

t, par suite, le vecteur

$$[X, Y] = \left.\frac{d\beta(t)}{dt}\right|_{t=0}$$

(cf. proposition 2 de la leçon 4) appartient à $\mathfrak{h}$. Donc, $\mathfrak{h}$ est une sous-algèbre de l'algèbre de Lie $\mathfrak{g}$. Pour achever la démonstration, il reste à remarquer que les groupuscules H et $E\mathfrak{h}$ sont manifestement équivalents. □

On voit, en particulier, que pour un sous-groupuscule H localement plan, la sous-algèbre $\mathfrak{h} \subset \mathfrak{l}(G)$ est composée de tous les éléments $X \in \mathfrak{l}(G)$ tels que exp $tX \in H$ pour tout t (avec $|t|$ assez petit). Cette sous-algèbre s'identifie de façon naturelle à l'algèbre de Lie $\mathfrak{l}(H)$ du sous-groupuscule H (traité comme un sous-groupuscule de Lie): on la désignera donc par $\mathfrak{l}(H)$.

Ainsi, *pour tout groupuscule de Lie G, il existe entre l'ensemble de toutes les sous-algèbres de l'algèbre de Lie* $\mathfrak{g} = \mathfrak{l}(G)$ *et l'ensemble de tous les sous-groupuscules localement plans du groupuscule G une correspondance biunivoque qui à la sous-algèbre $\mathfrak{h} \subset \mathfrak{g}$ associe le sous-groupuscule* exp $\mathfrak{h}$ *et au sous-groupuscule $H \subset G$, la sous-algèbre* $\mathfrak{l}(H)$.

Cette correspondance sera appelée *correspondance de Lie*.

Fait remarquable, E. Cartan a démontré que la condition de planation locale des sous-groupuscules est superflue dans cette proposition

Proposition 1 (théorème de Cartan). *Tout sous-groupuscule H d'un groupuscule de Lie G est localement plan, donc dans l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$ il existe une sous-algèbre $\mathfrak{h}$ telle que*

$$H = \exp \mathfrak{h}.$$

En particulier, H est un groupuscule de Lie.

Démonstration. Soit $\mathfrak{h}$ l'ensemble des éléments $X \in \mathfrak{g}$ tels que $\exp tX \in H$ pour tout t ($|t|$ assez petit). Nous allons montrer que $\mathfrak{h}$ est une sous-algèbre de l'algèbre $\mathfrak{g}$ et que $\exp \mathfrak{h} = H$. Divisons la démonstration en plusieurs lemmes.

Lemme 1. *Si pour un élément $X \in \mathfrak{g}$, il existe une suite $\{X_i\}$ convergeant vers X de telle sorte que*

$$\exp t_i X_i \in H$$

pour des $t_i \in \mathbb{R}$ tendant vers 0 lorsque $i \to \infty$, alors $X \in \mathfrak{h}$.

Démonstration. Etant donné que $\prod \exp(-t_i X_i) = \prod \exp (t_i X_i)^{-1} \in H$, on peut sans perdre en généralité admettre que $t_i > 0$. Soit $t > 0$ un nombre tel que l'élément $\exp tX \in G$ soit défini, et supposons que

$$k_i(t) = \left[\frac{t}{t_i}\right] \quad \left(\text{partie entière de } \frac{t}{t_i}\right).$$

Comme $\frac{t}{t_i} - 1 < k_i(t) \leqslant \frac{t}{t_i}$, il vient $\lim\limits_{t\to\infty} t_i k_i(t) = t$, et, par suite,

$$\lim_{t\to\infty} \exp (t_i k_i(t) X_i) = \exp tX.$$

Or

$$\exp (t_i k_i(t) X_i) = (\exp t_i X_i)^{k_i(t)} \in H,$$

et le sous-groupuscule H est par hypothèse localement fermé, donc $\exp tX \in H$, et, par suite, $X \in \mathfrak{h}$. □

Lemme 2. *Le sous-ensemble $\mathfrak{h} \subset \mathfrak{g}$ est une sous-algèbre.*

Démonstration. Comme $\exp (t(sX)) = \exp ((ts) X)$, pour tout $X \in \mathfrak{h}$ et tout $s \in \mathbb{R}$, on a l'appartenance $sX \in \mathfrak{h}$ (car t peut toujours être choisi suffisamment petit).

Soient $X, Y \in \mathfrak{h}$. Alors $\exp \mathfrak{J}(tX, tY) = \exp tX \cdot \exp tY \in H$, c'est-à-dire que $\exp t(X + Y + Z_t) \in H$, où $Z_t = O(t)$. Choisissons une suite arbitraire $\{t_i\}$ convergeant vers zéro et posons $X_i = X + Y + Z_{t_i}$. La suite $\{X_i\}$ vérifie (par rapport à l'élément $X + Y$) toutes les conditions du lemme 1. Donc, d'après ce lemme, $X + Y \in \mathfrak{h}$.

De façon analogue, comme

$$\exp tX \cdot \exp tY \cdot (\exp tX)^{-1} (\exp tY)^{-1} = \exp t^2([X, Y] + O(t)) \in H,$$

il vient $[X, Y] \in \mathfrak{h}$. □

Soit maintenant $\mathfrak{k}$ un sous-espace supplémentaire de $\mathfrak{h}$ dans $\mathfrak{g}$, c'est-à-dire tel que

(2) $$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{k}.$$

Lemme 3. *Il existe dans $\mathfrak{k}$ un voisinage V de 0 tel que*

$$\exp Y \notin H$$

pour aucun élément Y non nul de $\mathfrak{k}$.

Démonstration. Munissons $\mathfrak{k}$ d'une norme $\| \ \|$ (par exemple, euclidienne) et considérons l'ensemble B des éléments $Y \in \mathfrak{k}$ tels que $1 \leqslant \| Y \| \leqslant 2$. Si le lemme 3 est faux, il existe dans l'algèbre $\mathfrak{g}$ une suite $Y_i \to 0$ telle que $Y_i \in \mathfrak{k}$ et $\exp Y_i \in H$. Choisissons des nombres entiers n_i tels que $X_i = n_i Y_i \in B$ (il est clair que c'est toujours possible). L'ensemble B étant compact, on peut sans nuire à la généralité admettre que la suite $\{X_i\}$ converge. Soit X sa limite. La suite $\{X_i\}$ remplissant toutes les conditions du lemme 1 (avec $t_i = 1/n_i$), l'élément X (qui est visiblement non nul) appartient à la sous-algèbre $\mathfrak{h}$. Or ceci est impossible, puisque $X \in \mathfrak{k}$ et $\mathfrak{h} \cap \mathfrak{k} = 0$. Donc, le lemme 3 est vraie. □

Lemme 4. *Dans le groupuscule de Lie G on a*

$$H = \exp \mathfrak{h}.$$

Démonstration. Comme $\exp \mathfrak{h} \subset H$ par construction, il suffit simplement de prouver que $\exp \mathfrak{h} \supset H$. Sans perdre en généralité on peut admettre que le groupuscule de Lie G est un voisinage canonique de l'unité correspondant à la décomposition (2) de l'algèbre $\mathfrak{g}$ (cf. définition 4 de la leçon 5). Bien plus, on peut admettre que G est l'image par l'application

$$X + Y \mapsto \exp X \cdot \exp Y, \quad X \in \mathfrak{h}, \quad Y \subset \mathfrak{k},$$

d'un voisinage du zéro de $\mathfrak{g}$ de la forme $U \oplus V$, où U est un voisinage du zéro de $\mathfrak{h}$ et V un voisinage du zéro de $\mathfrak{k}$, tel qu'il figure dans le lemme 3. Or, dans ces conditions, l'inclusion $\exp \mathfrak{h} \supset H$ est évidente, puisque si $\exp X \cdot \exp Y \in H$, alors $\exp Y \in H$ (car $\exp X \in H$), et, par suite, $Y = 0$ en vertu du lemme 3. □

Ceci achève la démonstration de la proposition 1. □

Corollaire. *La correspondance de Lie est une correspondance biunivoque canonique entre l'ensemble des sous-groupuscules du groupuscule de Lie G et l'ensemble des sous-algèbres de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$.* □

En remarquant que ces deux ensembles sont des treillis pour l'inclusion, on obtient immédiatement que *la correspondance de Lie est un isomorphisme de ces espaces.*

Définition 2. On dit qu'un sous-groupuscule H d'un groupuscule de Lie G est *invariant* (ou *distingué*, ou *normal*) si $aba^{-1} \in H$ pour tous éléments $a \in G$, $b \in H$ tels que l'élément aba^{-1} soit défini.

Définition 3. On dit qu'une sous-algèbre $\mathfrak{h}$ d'une algèbre de Lie $\mathfrak{g}$ est un *idéal* si $[X, Y] \in \mathfrak{h}$ pour tous éléments $X \in \mathfrak{g}$, $Y \in \mathfrak{h}$.

Proposition 2. *La correspondance de Lie associe aux sous-groupuscules invariants d'un groupuscule de Lie G des idéaux de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$, et aux idéaux, des sous-groupuscules invariants.*

Démonstration. Soient $\mathfrak{h}$ un idéal d'une algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$, $X \in \mathfrak{g}$, $Y \in \mathfrak{h}$. De $[X, Y] \in \mathfrak{h}$, il s'ensuit immédiatement que $\mathfrak{H}_n(X, Y) \in \mathfrak{h}$ pour tout $n > 1$, et, par suite,

$$\mathfrak{H}(X, Y) = X + Y^*,$$

où $Y^* = Y + \sum_{n>1} \mathfrak{H}_n(X, Y) \in \mathfrak{h}$ (si bien sûr la série $\mathfrak{H}(X, Y)$ converge). Comme $[X + Y^*, -X] = [X, Y^*] \in \mathfrak{h}$ et donc que $\mathfrak{H}_n(X + Y^*, -X) \in \mathfrak{h}$ pour $n > 1$, il s'ensuit de là que

$$\mathfrak{H}(\mathfrak{H}(X, Y), -X) = X + Y^* - X + \sum_{n>1} \mathfrak{H}_n(X + Y^*, -X) \in \mathfrak{h},$$

et, par suite,

$$\exp X \cdot \exp Y \cdot (\exp X)^{-1} = \exp \mathfrak{H}(\mathfrak{H}(X, Y), -X) \in \exp \mathfrak{h}.$$

Par conséquent, le sous-groupuscule $\exp \mathfrak{h}$ est invariant.

Réciproquement, soit $H = \exp \mathfrak{h}$ un sous-groupuscule invariant du groupuscule de Lie G. Alors, pour tous éléments $X \in \mathfrak{g}$, $Y \in \mathfrak{h}$ et t ($|t|$ est assez petit), on aura

$$\exp tX \cdot \exp tY \cdot (\exp tX)^{-1} \in H,$$

et, par suite,

$$\exp tX \cdot \exp tY \cdot (\exp tX)^{-1} \cdot (\exp tY)^{-1} = \\ = \exp t^2 ([X, Y] + O(t)) \in H,$$

d'où au moyen de raisonnements déjà connus (cf. démonstration du lemme 2), on déduit que $[X, Y] \in \mathfrak{h}$. Donc, la sous-algèbre $\mathfrak{h} = \mathfrak{l}(H)$ est un idéal. □

Soit H un sous-groupuscule d'un groupuscule de Lie G. On dira que des éléments $a, b \in G$ sont *congrus modulo H* si $a^{-1}b \in H$. Il est clair que sur un voisinage assez petit de l'unité du groupuscule G, cette relation est une relation d'équivalence dont les classes d'équivalence sont les classes aH (plus exactement, leurs intersections avec un voisinage de l'unité). Soit G/H l'ensemble de ces classes.

En décomposant l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$ en une somme directe $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{k}$ de l'algèbre de Lie $\mathfrak{h}$ de H et d'un sous-espace $\mathfrak{k}$

et en considérant les coordonnées canoniques $x^1, \ldots, x^n$ définies par cette décomposition, on trouve immédiatement que les classes $aH \in G/H$ sont définies dans les coordonnées $x^1, \ldots, x^n$ par des équations de la forme

$$(3)\qquad x^{m+1} = a^1, \ldots, x^n = a^{n-m} \qquad (n = \dim G,\ m = \dim H),$$

où $a^1, \ldots, a^{n-m}$ sont des nombres réels quelconques (assez petits en module). Ceci montre qu'en associant à la classe aH le point $(a^1, \ldots, a^{n-m}) \in \mathbb{R}^{n-m}$, on obtient une application bijective de l'ensemble G/H sur un ouvert de l'espace $\mathbb{R}^{n-m}$, c'est-à-dire qu'on obtient une carte. Etant donné que pour tout autre choix de $\mathfrak{f}$ on obtient, comme il est aisé de le voir, une carte compatible, ceci munit l'ensemble G/H d'une structure topologique et d'une structure différentiable qui en font une variété différentiable (qui est difféomorphe à un ouvert de l'espace $\mathbb{R}^{n-m}$ et donc est de dimension $n - m$).

Si le sous-groupuscule H est invariant, la formule $aH \cdot bH = abH$ définit de façon unique au voisinage du point $H = eH$ de la variété G/H une multiplication pour laquelle elle est visiblement un groupuscule de Lie d'unité H.

Définition 4. Le groupuscule de Lie G/H s'appelle *groupuscule quotient* du groupuscule de Lie G par le sous-groupuscule invariant H.

Soit maintenant $\mathfrak{h}$ un idéal d'une algèbre de Lie $\mathfrak{g}$. Considérons l'espace quotient $\mathfrak{g}/\mathfrak{h}$ dont les éléments sont les classes $x + \mathfrak{h}$, $x \in \mathfrak{g}$. Il est évident que la formule

$$[x + \mathfrak{h}, y + \mathfrak{h}] = [x, y] + \mathfrak{h}$$

définit de façon unique l'opération $[\ ,\ ]$ sur l'espace quotient $\mathfrak{g}/\mathfrak{h}$ et en fait une algèbre de Lie.

Définition 5. L'algèbre de Lie $\mathfrak{g}/\mathfrak{h}$ s'appelle *algèbre quotient* de l'algèbre de Lie $\mathfrak{g}$ par l'idéal $\mathfrak{h}$.

En particulier, pour tout sous-groupuscule invariant H du groupuscule de Lie G est définie l'algèbre quotient $\mathfrak{l}(G)/\mathfrak{l}(H)$ de l'algèbre de Lie $\mathfrak{l}(G)$ par son idéal $\mathfrak{l}(H)$. Comparons cette algèbre quotient à l'algèbre de Lie $\mathfrak{l}(G/H)$ du groupuscule quotient G/H.

Proposition 3. *Les algèbres de Lie $\mathfrak{l}(G/H)$ et $\mathfrak{l}(G)/\mathfrak{l}(H)$ sont canoniquement isomorphes:*

$$\mathfrak{l}(G/H) \approx \mathfrak{l}(G)/\mathfrak{l}(H).$$

Démonstration. En appliquant le foncteur E: ALG$_f$-LIE $\to$ GR-LOC aux algèbres de Lie $\mathfrak{l}(G/H)$ et $\mathfrak{l}(G)/\mathfrak{l}(H)$, on obtient, d'une part, un groupuscule de Lie $E(\mathfrak{l}(G/H))$ qui est canoniquement isomorphe au groupuscule G/H et, d'autre part, le groupuscule de Lie $E(\mathfrak{l}(G)/\mathfrak{l}(H))$. Donc, la proposition 3 revient

à affirmer l'existence de l'isomorphisme

$$G/H \approx E\,(\mathfrak{l}(G)/\mathfrak{l}(H)).$$

C'est sous cette forme que nous allons la démontrer.

Soient $\mathfrak{g} = \mathfrak{l}(G)$ et $\mathfrak{h} = \mathfrak{l}(H)$. Les éléments de l'algèbre quotient $\mathfrak{g}/\mathfrak{h}$, c'est-à-dire les classes $X + \mathfrak{h}$ de l'algèbre $\mathfrak{g}$ suivant l'idéal $\mathfrak{h}$ sont les éléments du groupuscule $E(\mathfrak{l}(G)/\mathfrak{l}(H)) = E(\mathfrak{g}/\mathfrak{h})$. Montrons que l'application exponentielle $\exp: \mathfrak{g} \to G$ associe à toute classe $X + \mathfrak{h}$ une classe aH suivant $H = \exp \mathfrak{h}$. En effet, on sait déjà que pour tous éléments $X \in \mathfrak{g}$ et $Y \in \mathfrak{h}$ assez proches de 0, on a

$$\mathfrak{H}(X, Y) = X + Y^*,$$

où $Y^* \in \mathfrak{h}$. Ceci étant,

$$Y^* = Y + \sum_{n>1} \mathfrak{H}_n(X, Y),$$

d'où il s'ensuit immédiatement que pour tout X fixe, l'application $Y \to Y^*$ est bijective (dans un voisinage de 0). Donc, l'application

$$X + Y^* \mapsto \exp(X + Y^*) = \exp \mathfrak{H}(X, Y) = \exp X \cdot \exp Y$$

associe à toute classe $X + \mathfrak{h}$ une classe $\exp X \cdot H$.

Cette application de la variété $E(\mathfrak{g}/\mathfrak{h})$ sur la variété G/H est visiblement un difféomorphisme en l'unité. Donc, pour achever la démonstration de la proposition 3, il nous reste simplement à montrer qu'elle transforme un produit en un produit. Or, ceci est évident, puisque pour tous éléments $X_1, X_2 \in \mathfrak{g}$ et $Y_1, Y_2 \in \mathfrak{h}$, on a

$$(X_1 + Y_1^*)\cdot(X_2 + Y_2^*) = \mathfrak{H}(X_1 + Y_1^*, X_2 + Y_2^*) =$$
$$= \mathfrak{H}(X_1, X_2) + Y^*,$$

où $Y^* \in \mathfrak{h}$, et, par suite,

$$(X_1 + Y_1^*)\cdot(X_2 + Y_2^*) \mapsto \exp \mathfrak{H}(X_1, X_2)\cdot\exp Y^* \in a_1H\cdot a_2H,$$

où $a_1 = \exp X_1$, $a_2 = \exp X_2$. □

Récapitulons les résultats acquis dans le théorème suivant:

Théorème 1. *Pour tout groupuscule de Lie G, la correspondance de Lie est une application bijective canonique du treillis des sous-groupuscules du groupuscule de Lie G sur le treillis des sous-algèbres de l'algèbre de Lie* $\mathfrak{g} = \mathfrak{l}(G)$.

Cette application transforme les sous-groupuscules invariants en les idéaux de l'algèbre $\mathfrak{g}$ *et, réciproquement, les idéaux en les sous-groupuscules invariants.*

L'algèbre de Lie de tout groupuscule quotient du groupuscule G par un sous-groupuscule invariant H est l'algèbre quotient de l'algèbre de Lie $\mathfrak{l}(G)$ *par l'idéal correspondant.* □

Nous avons signalé à maintes occasions que tous les résultats des dernières leçons concernaient des groupuscules a n a l y t i q u e s (c'est-à-dire de classe C^ω). Il se trouve que nous n'avons pratiquement rien perdu en généralité. Pour formuler exactement cette affirmation, nous fixerons une classe de différentiabilité C^r ($r \geqslant 2$) et nous appellerons *groupuscules différentiables* les groupuscules qui sont des variétés de classe C^r, et *groupuscules analytiques*, les groupuscules de classe C^ω. On aura alors le théorème suivant qui montre qu'en théorie des groupuscules de Lie on ne perd rien en généralité en se restreignant aux groupuscules analytiques:

Théorème 2. *Tout groupuscule différentiable est isomorphe* (*dans la catégorie des groupuscules différentiables*) *à un groupuscule analytique.*

La d é m o n s t r a t i o n de ce théorème est basée sur une propriété intéressante en soi du foncteur de Lie que nous allons discuter préalablement.

Soient C et D deux catégories, $F: \mathsf{C} \to \mathsf{D}$ un foncteur quelconque. Pour tout couple A, B d'objets de la catégorie C, le foncteur F définit une application

$$\mathrm{Mor}_{\mathsf{C}}(A, B) \to \mathrm{Mor}_{\mathsf{D}}(FA, FB)$$

de l'ensemble des morphismes $A \to B$ de la catégorie C dans l'ensemble des morphismes $FA \to FB$ de la catégorie D. On dit que le foncteur F est *fidèle* si cette application est bijective quels que soient A et B, c'est-à-dire si pour tout morphisme $\beta: FA \to FB$ il existe un morphisme $\alpha: A \to B$ et un seul tel que $F\alpha = \beta$.

Il est évident que si le foncteur F est fidèle, les objets A et B de la catégorie C sont isomorphes si et seulement si le sont les objets FA et FB de la catégorie D.

Dans la proposition suivante on comprendra par GR-LOC la catégorie des groupuscules de Lie différentiables.

Proposition 4. *Le foncteur de Lie* GR-LOC $\to$ ALG_f-LIE *est fidèle.*

En particulier, des groupuscules de Lie sont isomorphes si et seulement si leurs algèbres de Lie le sont.

Le théorème 2 résulte immédiatement de cette proposition. En effet, soient G un groupuscule différentiable, $\mathfrak{g} = \mathfrak{l}(G)$ son algèbre de Lie. Le foncteur E construit dans la leçon précédente associe à l'algèbre de Lie $\mathfrak{g}$ un groupuscule analytique $E\mathfrak{g}$ dont l'algèbre de Lie $\mathfrak{l}(E\mathfrak{g})$ est isomorphe à l'algèbre $\mathfrak{g}$. Donc, ce groupuscule est isomorphe au groupuscule G. □

A noter que la proposition 4 est trivialement évidente pour les groupuscules analytiques, puisque *tout foncteur quasiréciproque F*:

$\mathbf{C} \to \mathbf{D}$ *est fidèle*. En effet, si $G: \mathbf{D} \to \mathbf{C}$ est un foncteur quasiréciproque, tout morphisme $\alpha: A \to B$ est justiciable du diagramme commutatif

$$\begin{array}{ccc} A & \xrightarrow{\varepsilon_A} & GFA \\ \downarrow & & \downarrow GF\alpha \\ B & \xrightarrow{\varepsilon_B} & GFB \end{array}$$

dont les flèches horizontales figurent des isomorphismes. Donc, si $F\alpha = F\beta$, alors

$$\alpha = \varepsilon_B^{-1} \circ GF\alpha \circ \varepsilon_A = \varepsilon_B^{-1} \circ GF\beta \circ \varepsilon_A = \beta.$$

Par conséquent, si pour un morphisme $\gamma: FA \to FB$, il existe un morphisme $\alpha: A \to B$ tel que $F\alpha = \gamma$, alors ce morphisme est unique. On démontre de façon analogue que pour un morphisme $\alpha: GS \to GT$, où S et T sont des objets de la catégorie $\mathbf{D}$, il ne peut exister qu'un seul morphisme $\gamma: S \to T$ tel que $G\gamma = \alpha$. Par ailleurs, pour tout morphisme $\gamma: FA \to FB$, le morphisme $\alpha = \varepsilon_B^{-1} \circ G\gamma \circ \varepsilon_A$ possède la propriété suivante: $GF\alpha = \varepsilon_B \circ \alpha \circ \varepsilon_A^{-1} = G\gamma$. Donc, $F\alpha = \gamma$. □

La démonstration de la proposition 4 est basée sur la théorie des équations aux différentielles totales (équations de Pfaff). Commençons par exposer cette théorie sous la forme intrinsèque (sans utiliser les coordonnées).

Soient P et Q des variétés différentiables.

Définition 6. On appelle *système de Pfaff* de P dans Q une fonction

$$f: (p, q) \mapsto f(p, q),$$

associant à tout point $(p, q) \in P \times Q$ une application linéaire

$$f(p, q): \mathsf{T}_p(P) \to \mathsf{T}_q(Q),$$

dépendant différentiablement de p et de q.

Si $x^1, \ldots, x^n$ et $y^1, \ldots, y^m$ sont des coordonnées locales respectivement dans les variétés P et Q, alors dans les coordonnées associées sur les espaces vectoriels $\mathsf{T}_p(P)$ et $\mathsf{T}_q(Q)$ l'application $f(p, q)$ sera définie par une $n \times m$-matrice dont les éléments sont des fonctions des coordonnées $x^1, \ldots, x^n$ et $y^1, \ldots, y^m$. La condition de dépendance différentiable de l'application $f(p, q)$ par rapport à p et q signifie que ces fonctions sont différentiables.

Remarque 1. En termes plus intrinsèques, la notion de système de Pfaff se définit à l'aide de la notion de *fibré induit*. Nous ne rap-

pellerons pas cette notion (car nous n'en aurons besoin nulle part ailleurs) et remarquerons seulement qu'aux termes de la définition 6, le système de Pfaff n'est autre chose que l'application sur $P \times Q$ du fibré induit du fibré tangent $\mathsf{T}(P)$ par la projection $P \times Q \to P$, dans le fibré induit du fibré tangent $\mathsf{T}(Q)$ par la projection $P \times Q \to Q$. Cette approche nous dispense d'expliquer la condition de dépendance différentiable des applications $f(p, q)$ par rapport à p et q.

Définition 7. On appelle *intégrale* d'un système de Pfaff f sur un ouvert $U \subset P$ une application $\varphi: U \to Q$ telle que

$$f(u, \varphi x) = (d\varphi)_u$$

pour tout point $u \in U$. Un système de Pfaff f est *intégrable* si pour tout point $(p_0, q_0) \in P \times Q$, il existe une intégrale $\varphi: U \to Q$ du système f, définie sur un voisinage U de p_0 dans P, telle que $\varphi(p_0) = q_0$.

Lemme 1. *Deux intégrales quelconques $\varphi: U \to Q$ et $\varphi': U' \to Q$ définies sur des voisinages U et U' d'un point p_0 et telles que $\varphi(p_0) = \varphi'(p_0)$ sont confondues sur un voisinage du point p_0.*

Nous prouverons ce lemme plus bas.

Pour obtenir des conditions commodes d'intégrabilité d'un système de Pfaff (et prouver par la même occasion le lemme 1), il faut généraliser un peu le problème.

Soient M une variété différentiable, $\pi: \mathsf{T}(M) \to M$ son fibré tangent.

Définition 8. Un fibré vectoriel $\pi_1: E \to M$ est un *sous-fibré* d'un fibré tangent $\pi: \mathsf{T}(M) \to M$ si $E \subset \mathsf{T}(M)$, l'injection $E \to \mathsf{T}(M)$ est différentiable et $\pi_1 = \pi|_E$. Tout sous-fibré est défini de façon unique par la donnée de la variété E et en principe nous l'identifierons à E. On notera E_a la fibre $\pi_1^{-1}(a)$ du fibré E au-dessus du point $a \in M$. Cette fibre est le sous-espace $\mathsf{T}_a(M) \cap E$ de l'espace tangent $\mathsf{T}_a(M)$ dont la dimension est la même pour tous les a. Cette dimension sera désignée par m.

Pour tout ensemble ouvert $U \subset M$ est défini un fibré vectoriel $\pi_1^{-1}(U) \to U$ des mêmes fibres que le fibré E. C'est un sous-fibré du fibré tangent $\mathsf{T}(U) \to U$. Nous le désignerons par $E|_U$ et l'appellerons *restriction* du sous-fibré E à U.

D'après cette définition $\mathsf{T}(M)|_U = \mathsf{T}(U)$.

Soient W une variété de dimension m, w_0 un point de W. En général, la variété W sera un voisinage du point w_0 dans une certaine variété (par exemple, $\mathbb{R}^m$), mais en principe W peut être arbitraire.

Définition 9. On dit qu'une application différentiable $\Phi: W \to M$ est *intégrale relativement à un sous-fibré* $E \subset \mathsf{T}(M)$ si en tout point $w \in W$ sa différentielle

$$(d\Phi)_w: \mathsf{T}_w(W) \to \mathsf{T}_a(M), \qquad a = \Phi w,$$

est un monomorphisme sur la fibre E_a du sous-fibré E.

Conformément aux notations en usage dans la théorie des ensembles, on écrira $\Phi: (W, w_0) \to (M, a_0)$ si $a_0 = \Phi w_0$.

Définition 10. Un sous-fibré E d'un fibré tangent $\mathsf{T}(M)$ est *intégrable* si pour tout point $a_0 \in M$, il existe une application $\Phi: (W, w_0) \to (M, a_0)$ intégrale relativement à E.

Lemme 2. *Si un sous-fibré E est intégrable, alors pour tout point $a_0 \in M$ et tout couple $\Phi: (W, w_0) \to (M, a_0)$ et $\Phi': (W', w_0') \to (M, a_0)$ d'applications intégrales relativement à E, il existe dans les variétés W et W' des voisinages V et V' des points w_0 et w_0' et un difféomorphisme $\beta: V' \to V$ tels que $\Phi' = \Phi \circ \beta$ sur V'.*

Nous prouverons ce lemme ultérieurement.

Soient f un système de Pfaff sur P dans Q et $M = P \times Q$. Pour tout point $a = (p, q) \in M$, considérons le *graphe* de l'application $f(p, q): \mathsf{T}_p(P) \to \mathsf{T}_q(Q)$, c'est-à-dire le sous-ensemble E_a de la somme directe

$$\mathsf{T}_a(M) = \mathsf{T}_p(P) \oplus \mathsf{T}_q(Q),$$

formé des vecteurs (A, B), $A \in \mathsf{T}_p(P)$, $B \in \mathsf{T}_q(Q)$ tels que

$$B = f(p, q)A.$$

L'application $f(p, q)$ étant linéaire, ce sous-ensemble est un sous-espace.

Posons

$$E = \bigcup_{a \in M} E_a.$$

L'ensemble E est un sous-ensemble de $\mathsf{T}(M)$ et la restriction $\pi_1: E \to M$ de la projection $\pi: \mathsf{T}(M) \to M$ à E est telle que $\pi_1^{-1}(a) = E_a$ pour tout point $a \in M$.

Si $(U \times V, h \times k)$ est une carte de M qui soit le produit des cartes (U, k) et (V, k) de P et Q respectivement, et si

$$W = \pi_1^{-1}(U \times V) = \bigcup_{a \in U \times V} E_a,$$

alors l'application

$$W \to (U \times V) \times \mathbb{R}^m, \quad m = \dim P,$$

définie par la formule

$$(A, B) \mapsto (\pi(A, B), \dot{h}A),$$

où $\dot{h}$: $\mathsf{T}_p(P) \to \mathbb{R}^m$, $p = \pi A$, est la carte correspondant à la carte h, est visiblement bijective. L'application réciproque H associe au point (a, x), où $a = (p, q) \in U \times V$, et $x \in \mathbb{R}^m$, le vecteur $(\dot{h}^{-1}x, f(p, q)\dot{h}^{-1}x)$, et, par suite, est telle que pour tout point $a \in U \times V$ l'application $x \mapsto H(a, x)$ est un isomorphisme de $\mathbb{R}^m$ sur l'espace vectoriel E_a.

Sous ces conditions, si (U', h') est une autre carte de P, et

$$H' : (U' \times V) \times \mathbb{R}^m \to W' = \bigcup_{a \in U' \times V} E_a$$

l'application correspondante, alors (sous réserve que $U \cap U' \neq \varnothing$) pour tout point $a \in (U \times V) \cap (U' \times V)$ la composée de l'application $x \mapsto H(a, x)$ et de l'application réciproque de $x \mapsto H'(a, x)$ sera visiblement confondue avec l'application $\dot{h}' \circ \dot{h}^{-1}$: $\mathbb{R}^m \to \mathbb{R}^m$, et par suite, dépendra différentiablement du point a.

Tout cela signifie que (E, π_1, M) est un fibré vectoriel. Par ailleurs, il est clair que l'injection $E \to \mathsf{T}(M)$ est différentiable, et, par suite (puisque $\pi_1 = \pi|_E$ par construction), ce fibré est un sous-fibré du fibré tangent $\mathsf{T}(M)$.

Définition 11. Le sous-fibré construit s'appelle *graphe* du système de Pfaff f.

Il est évident qu'une application φ: $U \to Q$ est une intégrale du système de Pfaff f si et seulement si l'application Φ: $U \to P \times Q$ définie par la formule

$$\Phi u = (u, \varphi u), \qquad u \in U, \tag{4}$$

est intégrale relativement au graphe E du système f. D'autre part, si Ψ: $W \to P \times Q$ est une application intégrale relativement au graphe E, et $\Psi w = (\alpha w, \psi w)$, où α: $W \to P$, ψ: $W \to Q$, alors pour tout point $w_0 \in W$, la différentielle $(d\alpha)_{w_0}$ de l'application α sera, ce qui est aisé à voir, un isomorphisme et par suite — en passant éventuellement de W à un voisinage plus petit du point w_0 — l'application α est elle-même un difféomorphisme de la variété W sur un voisinage U du point $u_0 = \alpha w_0$. Sous ces conditions, l'application $\Phi = \Psi \circ \alpha^{-1}$: $U \to P \times Q$ qui est toujours intégrale relativement au sous-fibré E, sera définie par la formule (4) (avec $\varphi = \psi \circ \alpha^{-1}$) et, par conséquent, déterminera l'intégrale φ du système de Pfaff f. Ceci prouve que *le système de Pfaff sur P dans Q est intégrable si et seulement si son graphe est un sous-fibré intégrable du fibré tangent de la variété $P \times Q$.*

Nous sommes en mesure de prouver maintenant le lemme 1 (en admettant la validité du lemme 2).

D é m o n s t r a t i o n d u l e m m e 1. Soient φ: $(U, u_0) \to (Q, q_0)$ et φ': $(U', u_0) \to (Q, q_0)$ deux intégrales d'un système

de Pfaff f, confondues en un point $u_0 \in U \cap U'$, $\Phi: u \mapsto (u, \varphi u)$ et $\Phi': u \mapsto (u, \varphi' u)$ les applications correspondantes (4), intégrales relativement au graphe E de ce système. D'après le lemme 2, il existe des voisinages V et V' du point u_0 et un difféomorphisme β: $V' \to V$ tels que $\Phi' = \Phi \circ \beta$ sur V'. Par projection sur U, on trouve, en particulier, que $\beta u = u$ pour tout point $u \in V'$, c'est-à-dire que $V' = V$ et $\beta = \mathrm{id}$. Mais alors $\varphi' = \varphi' \circ \beta = \varphi'$ sur $V' = V$, ce qui prouve le lemme 1. □

Soit E un sous-fibré du fibré $\mathsf{T}(M)$. On dira qu'un champ de vecteurs $X \in \mathfrak{a}(M)$ *est situé dans* E si $X_a \in E_a$ pour tout point $a \in M$. Il est clair que de tels champs forment un sous-module du $\mathcal{F}(M)$-module $\mathfrak{a}(M)$. On notera ce sous-module $\mathfrak{a}(E)$.

Si le fibré E est trivial (est isomorphe au fibré $M \times \mathbb{R}^n \to M$), il existe sur M des champs de vecteurs $X_1, \ldots, X_m$ tels que pour tout point $a \in M$, les vecteurs $(X_1)_a, \ldots, (X_m)_a$ forment une base de l'espace E_a, et, par suite, les champs $X_1, \ldots, X_m$ forment une base dans le $\mathcal{F}(M)$-module $\mathfrak{a}(E)$, c'est-à-dire que ce module est libre. De là il s'ensuit, en vertu de la trivialité locale du fibré E, que *pour tout sous-fibré E du fibré tangent $\mathsf{T}(M)$, il existe un recouvrement ouvert $\{U\}$ de la variété M, tel que pour tout élément U de ce recouvrement le $\mathcal{F}(U)$-module $\mathfrak{a}(E|_U)$ est libre.*

Définition 12. Un sous-fibré E est dit *involutif* si le sous-module $\mathfrak{a}(E)$ est une sous-algèbre de l'algèbre de Lie $\mathfrak{a}(M)$, c'est-à-dire si $[X, Y] \in \mathfrak{a}(E)$ pour tous champs $X, Y \in \mathfrak{a}(E)$.

Plus bas, nous aurons affaire au cas où le sous-module $\mathfrak{a}(E)$ est engendré par une sous-algèbre $\mathfrak{g}$ (de dimension finie) de l'algèbre de Lie $\mathfrak{a}(M)$, c'est-à-dire que tout élément de $\mathfrak{a}(E)$ est une combinaison linéaire de champs de $\mathfrak{g}$ dont les coefficients appartiennent à $\mathcal{F}(M)$ (conformément aux notations classiques de la théorie des modules, on désignera un tel sous-module par $\mathcal{F}(M)\mathfrak{g}$). Il est immédiat de voir que dans ce cas la condition d'involutivité est remplie, c'est-à-dire que *tout sous-module engendré par une sous-algèbre est lui-même une sous-algèbre.* En effet, il suffit de toute évidence de montrer que pour tous champs $X, Y \in \mathfrak{g}$ et toute fonction $f \in \mathcal{F}(M)$, le champ $[fX, Y]$ est situé dans $\mathcal{F}(M)\mathfrak{g}$. Mais comme Y est une dérivation, $Y \circ fX = Yf \cdot X + f \cdot Y \circ X$ et, par suite,

$$[fX, Y] = fX \circ Y - Y \circ fX = f \cdot (X \circ Y - Y \circ X) - Yf \cdot X =$$
$$= f \cdot [X, Y] - Yf \cdot X \in \mathcal{F}(M)\,\mathfrak{g},$$

car $[X, Y] \in \mathfrak{g}$ et $Yf \in \mathcal{F}(M)$. □

Proposition 5 (théorème de Frobenius). *Un sous-fibré E est intégrable si et seulement s'il est involutif.*

Avant de prouver cette proposition, nous allons nous en servir pour démontrer la proposition 4.

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 4. Soient G et H deux groupuscules de Lie différentiables, $f: \mathfrak{g} \to \mathfrak{h}$ un homomorphisme de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$ dans l'algèbre de Lie $\mathfrak{h} = \mathfrak{l}(H)$. Il nous faut montrer qu'il existe un homomorphisme $\varphi: G \to H$ pour lequel $\mathfrak{l}(\varphi) = f$ et que cet homomorphisme est unique.

A cet effet, en traitant f comme une application $\mathsf{T}_e(G) \to \mathsf{T}_e(H)$, définissons pour deux éléments quelconques $a \in G$ et $b \in H$ une application linéaire $f(a, b): \mathsf{T}_a(G) \to \mathsf{T}_b(H)$ en posant

$$f(a, b) = (dL_b)_e \circ f \circ (dL_a)_e^{-1}, \tag{5}$$

où comme toujours L_a et L_b sont des translations à gauche. Il est évident que les applications $f(a, b)$ dépendent différentiablement de a et de b, c'est-à-dire forment un système de Pfaff sur G dans H.

Pour tout homomorphisme $\varphi: G \to H$ de groupuscules de Lie et tout élément $a \in G$, on a la relation $\varphi \circ L_a = L_{\varphi a} \circ \varphi$ (qui signifie tout simplement que $\varphi(ax) = \varphi a\, \varphi x$), c'est-à-dire que $\varphi = L_{\varphi a} \circ \varphi \circ L_a^{-1}$. Donc,

$$(d\varphi)_a = (dL_{\varphi a})_e \circ (d\varphi)_e \circ (dL_a)_e^{-1}.$$

d'où il s'ensuit que si $f = \mathfrak{l}(\varphi)$, c'est-à-dire que $f = (d\varphi)_e$, alors

$$(d\varphi)_a = f(a, \varphi a),$$

c'est-à-dire que φ est une intégrale du système de Pfaff (5) telle que $\varphi e = e$.

Réciproquement, supposons que le système de Pfaff (5) possède une intégrale φ définie sur un voisinage de l'unité e du groupuscule G et telle que $\varphi e = e$. Puisque nous avons convenu d'identifier les groupuscules équivalents, nous pouvons, sans nuire à la généralité, admettre que l'intégrale φ est définie sur le groupuscule G tout entier.

Pour tout point fixe $a \in G$, l'application $\varphi \circ L_a: x \mapsto \varphi(ax)$ définie sur un voisinage du point e, vérifie la relation

$$\begin{aligned} d(\varphi \circ L_a)_x &= (d\varphi)_{ax} \circ (dL_a)_x = \\ &= (dL_{\varphi(ax)})_e \circ f \circ (dL_{ax})_e^{-1} \circ (dL_a)_x = \\ &= (dL_{\varphi(ax)})_e \circ f \circ (dL_x)_e^{-1} = f(x, \varphi(ax)), \end{aligned}$$

c'est-à-dire est une intégrale du système de Pfaff (5). De façon analogue, puisque

$$\begin{aligned} (dL_{\varphi a} \circ \varphi)_x &= (dL_{\varphi a})_{\varphi x} \circ (d\varphi)_x = \\ &= (dL_{\varphi a})_{\varphi x} \circ (dL_{\varphi x})_e \circ f \circ (dL_x)_e^{-1} = \\ &= (dL_{\varphi a\, \varphi x})_e \circ f \circ (dL_x)_e^{-1} = f(x, \varphi a\, \varphi x), \end{aligned}$$

l'application $L_{\varphi a} \circ \varphi : x \to \varphi a \, \varphi x$ est aussi une intégrale du système de Pfaff (5). Comme ces intégrales prennent la même valeur φa au point e, en vertu du lemme 1 elles sont confondues dans un voisinage de ce point. Donc, dans ce voisinage $\varphi(ax) = \varphi a \, \varphi x$, et, par suite, φ est un homomorphisme de G dans H. Comme

$$(d\varphi)_e = (dL_e)_e \circ f \circ (dL_e)_e^{-1} = f,$$

cet homomorphisme induit l'homomorphisme f d'algèbres de Lie.

Nous avons ainsi prouvé que les homomorphismes $\varphi : G \to H$ de groupuscules de Lie qui induisent l'homomorphisme $f : \mathfrak{g} \to \mathfrak{h}$ d'algèbres de Lie sont très exactement les intégrales du système de Pfaff (5) pour lesquelles $\varphi e = e$. Donc, pour prouver la proposition 4, il suffit de montrer que le système de Pfaff (5) est intégrable, c'est-à-dire qu'est intégrable le sous-fibré E du fibré tangent $\mathsf{T}(G \times H)$ qui est le graphe de ce système.

La fibre de ce graphe au-dessus du point $(e, e) \in G \times H$ est de toute évidence le graphe $\mathfrak{f}$ de l'homomorphisme f, c'est-à-dire le sous-espace de l'espace $\mathsf{T}_e(G \times H) = \mathsf{T}_e(G) \times \mathsf{T}_e(H)$, formé des couples de la forme (A, fA), où $A \in \mathsf{T}_e(G)$. On obtient la fibre du graphe au-dessus d'un point arbitraire $(a, b) \in G \times H$ en faisant agir l'opérateur linéaire $dL_{(a, b)} = dL_a \times dL_b$ sur $\mathfrak{f}$.

Si donc l'on considère sur $G \times H$ des champs de vecteurs invariants à gauche $X_1, \ldots, X_m$ dont les valeurs $X_1(e, e), \ldots, X_m(e, e)$ au point (e, e) forment une base du sous-espace $\mathfrak{f}$, alors leurs valeurs $X_1(a, b), \ldots, X_m(a, b)$ en un point quelconque $(a, b) \in G \times H$ formeront une base de la fibre du graphe E au-dessus du point (a, b). Ceci signifie que *les champs* $X_1, \ldots, X_m$ *forment une base du* $\mathcal{F}(G \times H)$*-module* $\mathfrak{a}(E)$ (de sorte que ce module est libre).

Pour reformuler cette affirmation en termes plus intrinsèques, on remarquera qu'étant le graphe d'un homomorphisme d'algèbres de Lie, le sous-espace $\mathfrak{f}$ de l'algèbre de Lie $\mathfrak{l}(G \times H) \subset \mathfrak{a}(G \times H)$ sera une sous-algèbre de cette algèbre, et, par suite, de l'algèbre de Lie $\mathfrak{a}(G \times H)$ tout entière. Si l'algèbre $\mathfrak{l}(G \times H)$ est traitée comme une algèbre de champs de vecteurs invariants à gauche, une base de $\mathfrak{f}$ sera constituée précisément des champs $X_1, \ldots, X_m$. Donc, le sous-module du $\mathcal{F}(G \times H)$-module $\mathfrak{a}(G \times H)$ de base $X_1, \ldots, X_m$ sera précisément le sous-module $\mathcal{F}(G \times H)\,\mathfrak{f}$ engendré par la sous-algèbre $\mathfrak{f}$. Or, on a remarqué plus haut qu'un sous-module engendré par une sous-algèbre est une sous-algèbre. Donc, le sous-module $\mathfrak{a}(E)$ est une sous-algèbre de l'algèbre $\mathfrak{a}(G \times H)$, si bien que le sous-fibré E est involutif. Il est donc intégrable d'après la proposition 5. □

Il ne nous reste plus qu'à démontrer la proposition 5 (et le lemme 2). La condition nécessaire de cette proposition s'établit sans peine:

Proposition 6. *Tout sous-fibré intégrable E est involutif.*

Démonstration. Soient $a_0 \in M$ et $\Phi: (W, w_0) \to (M, a_0)$ une application intégrale relativement à E. Comme l'application $(d\Phi)_w: \mathsf{T}_w(W) \to \mathsf{T}_{\Phi w}(M)$ est monomorphe en tout point $w \in W$, pour tout champ de vecteurs $X \in \mathfrak{a}(E)$, il existe sur W un seul champ X^Φ satisfaisant à la relation

$$X_{\Phi w} = (d\Phi)_w (X_w^\Phi),$$

c'est-à-dire Φ-lié au champ X. Sous ces conditions, pour tout couple de champs X, $Y \in \mathfrak{a}(E)$, les champs $[X, Y]$ et $[X^\Phi, Y^\Phi]$ sont aussi Φ-liés, c'est-à-dire qu'on a

$$[X, Y]_{\Phi w} = (d\Phi)_w [X^\Phi, Y^\Phi]_w.$$

Donc, $[X, Y]_{\Phi w} \in \operatorname{Im} (d\Phi)_w = E_{\Phi w}$. En particulier, $[X, Y]_{a_0} \in$ $\in E_{a_0}$. Comme le point $a_0 \in M$ est arbitraire, ceci prouve que $[X, Y] \in$ $\in \mathfrak{a}(E)$. Donc, le sous-fibré E est involutif. □

La démonstration de la réciproque est plus délicate.

Définition 13. On dit qu'un sous-fibré E d'un fibré tangent $\mathsf{T}(M)$ est *complètement intégrable* si la variété M est munie d'un atlas de cartes $(U, x^1, \ldots, x^n)$ telles que pour tout point $a \in U$ les m premiers vecteurs $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots, \left(\frac{\partial}{\partial x^m}\right)_a$ de la base $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots$ $\ldots, \left(\frac{\partial}{\partial x^n}\right)_a$ de l'espace $\mathsf{T}_a(M)$ forment une base de l'espace E_a, autrement dit ces vecteurs sont tels que les champs de vecteurs $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$ sur U forment une base du $\mathcal{F}(U)$-module $\mathfrak{a}(E|_U)$.

Il est immédiat de voir que *tout sous-fibré complètement intégrable est intégrable.* En effet, soit $a_0 \in M$ et supposons que $a_0 \in U$, où $(U, h) = (U, x^1, \ldots, x^n)$ est une carte de l'atlas de la définition 13. Soient par ailleurs $\mathcal{L}^m$ le plan $x^{m+1} = x^{m+1}(a_0), \ldots$ $\ldots, x^n = x^n(a_0)$ de l'espace $\mathbb{R}^n$ et $W = \mathcal{L}^m \cap h(U)$ (qui contient visiblement le point $w_0 = h(a_0)$). Soit enfin $\Phi: W \to M$ la restriction à W du difféomorphisme réciproque $h^{-1}: h(U) \to U$ (traité comme une application dans M). La différentielle $(d\Phi)_w$ de l'application Φ en un point $w \in W$ transforme la base canonique de l'espace $\mathsf{T}_w(W) = \mathbb{R}^m$ en les vecteurs $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots, \left(\frac{\partial}{\partial x^m}\right)_a$, où $a = \Phi w$, et, par suite, est un monomorphisme sur l'espace E_a.

L'application $\Phi: (W, w_0) \to (M, a_0)$ est donc une application intégrale relativement à E. Comme le point $a_0 \in M$ est arbitraire, ceci exprime que le sous-fibré E est intégrable. □

Bien plus, on démontre de façon analogue que *tout sous-fibré complètement intégrable E possède la propriété énoncée dans le lemme* 2, à savoir que pour tout couple $\Phi: (W, w_0) \to (M, a_0)$ et $\Phi': (W', w'_0 \to (M, a_0)$ d'applications intégrales relativement au sous-fibré E, on peut exhiber dans W et W' des voisinages V et V' de w_0 et w'_0 et un difféomorphisme $\beta: V' \to V$ tels que $\Phi' = \Phi \circ \beta$ sur V'. En effet, sans restreindre la généralité, on peut admettre que l'application Φ' est l'application construite ci-dessus au moyen de la carte $(U, x^1, \ldots, x^n)$ et que la variété W est un sous-ensemble ouvert de l'espace $\mathbb{R}^m$. Sous ces conditions, pour tout point $w = (w^1, \ldots, w^m) \in W$, les vecteurs $(d\Phi)_w\left(\frac{\partial}{\partial w^1}\right), \ldots, (d\Phi)_w\left(\frac{\partial}{\partial w^m}\right)$ formeront une base de l'espace E_a, $a = \Phi w$ et, par suite, s'exprimeront linéairement en fonction des vecteurs $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots, \left(\frac{\partial}{\partial x^m}\right)_a$. Par ailleurs, si $x^1(w), \ldots, x^n(w)$ sont les fonctions déterminant l'application Φ, alors

$$(d\Phi)_w\left(\frac{\partial}{\partial w^i}\right) = \left(\frac{\partial x^j}{\partial w^i}\right)_w \cdot \left(\frac{\partial}{\partial x^j}\right)_a, \quad i = 1, \ldots, m,$$

d'après la règle générale qui définit l'action de la différentielle d'une application différentiable. Donc, $\frac{\partial x^j}{\partial w^i} = 0$ sur W pour tout $i = 1, \ldots, m$ et tout $j = m + 1, \ldots, n$, et, par suite, $x^j(w) = \text{const}$ (plus exactement, $x^j(w) = x^j(a_0)$) pour $j = m + 1, \ldots, n$. Ceci signifie que le composé $\beta = h \circ \Phi$ de l'application Φ et du difféomorphisme de carte h est une application de W dans W'. L'application β étant visiblement étale, le point w_0 possède un voisinage V sur lequel elle est un difféomorphisme sur un voisinage V' du point w'_0. Pour achever la démonstration, il reste à remarquer que $\Phi = h^{-1} \circ (h \circ \Phi) = \Phi \circ \beta$ sur V. □

D'après ces remarques, pour prouver le lemme 2 et le reste de la proposition 5, il nous suffit de démontrer la proposition suivante:

Proposition 7. *Tout sous-fibré involutif E d'un fibré tangent* $\mathsf{T}(M)$ *est complètement intégrable.*

D é m o n s t r a t i o n. Soient $a_0 \in M$ et $(U, h) = (U, x^1, \ldots, x^n)$ une carte de M en a_0. Quitte à choisir un plus petit voisinage, on peut admettre que le $\mathcal{F}(U)$-module $\mathfrak{a}(E|_U)$ est libre, c'est-à-dire possède une base formée de m champs de vecteurs $X_1, \ldots, X_m$. Pour simplifier les formules, on peut admettre par ailleurs que $x^1(a_0) = 0, \ldots, x^n(a_0) = 0$.

Supposons d'abord que $m = 1$. Posons $X = X_1$ et considérons les composantes $X^1 = Xx^1, \ldots, X^n = Xx^n$ du champ de vecteurs X dans les coordonnées $x^1, \ldots, x^n$. Comme $X \neq 0$ dans U, on peut sans nuire à la généralité considérer que $X^n \neq 0$ dans U (il

suffit au besoin de changer le nom des coordonnées et de réduire U).

On sait que pour tout point $u \in U$, il existe une courbe intégrale $\varphi_u: I_u \to M$ du champ X, définie sur un intervalle I_u de $\mathbb{R}$ contenant 0, et telle que $\varphi_u(0) = u$. On peut admettre (quitte à réduire U) que l'intervalle I_u ne dépend pas de u (est le même intervalle I pour tous les u).

Considérons maintenant dans l'espace $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$, un ensemble ouvert $W \times I$, où W est l'ensemble de tous les points $w = (w^1, \ldots, w^{n-1}) \in \mathbb{R}^{n-1}$ pour lesquels $(w, 0) \in h(U)$, et l'application

$$\Phi: W \times I \to M$$

de cet ensemble dans une variété M définie par la formule

$$\Phi(w, t) = \varphi_u(t), \quad \text{où} \quad u = h^{-1}(w, 0).$$

Dans les coordonnées $w^1, \ldots, w^{n-1}, w^n = t$ et $x^1, \ldots, x^n$, cette application est déterminée par des fonctions $x^1(w, t), \ldots, x^n(w, t)$ telles que

$$\text{(6)} \qquad \frac{dx^i(w, t)}{dt} = X^i(x^1(w, t), \ldots, x^n(w, t)), \quad i = 1, \ldots, n,$$

et

$$x^1(w, 0) = w^1, \quad \ldots, \quad x^{n-1}(w, 0) = w^{n-1}, \quad x^n(w, 0) = 0$$

identiquement en w. On voit notamment que

$$\left(\frac{\partial x^i}{\partial w^j}\right)_{t=0} = \delta^i_j \text{ et } \left(\frac{\partial x^n}{\partial t}\right)_{t=0} = 0.$$

Combiné aux formules (6) ceci signifie qu'en tout point de la forme $(w, 0)$, et en particulier en $(0, 0)$, la matrice de la différentielle de l'application Φ est, dans les coordonnées $w^1, \ldots, w^{n-1}, w^n = t$ et $x^1, \ldots, x^n$, de la forme suivante

$$\begin{pmatrix} 1 & & & X^1 \\ & \ddots & 0 & \vdots \\ 0 & & 1 & X^{n-1} \\ 0 & \ldots & 0 & X^n \end{pmatrix},$$

donc n'est pas dégénérée (en vertu de la condition $X^n \neq 0$). Donc, l'application Φ est étale au point $0 = (0, 0) \in W \times I$, par conséquent, sans nuire à la généralité, on peut admettre que c'est un difféomorphisme de l'ensemble $W \times I$ sur le voisinage U. Le difféomorphisme réciproque Φ^{-1} définit sur U des coordonnées $w^1, \ldots$ $\ldots, w^{n-1}, w^n$ telles que visiblement $\frac{\partial}{\partial w^n} = X$ sur U.

Donc, dans un voisinage U du point a_0, il existe des coordonnées $w^1, \ldots, w^n$, telles que le champ $\frac{\partial}{\partial w^n}$ engendre le sous-module $\mathfrak{a}(E|_U)$. Comme le point a_0 a été arbitrairement choisi, ceci prouve la complète intégrabilité du sous-fibré E.

Ceci achève la démonstration de la proposition 6 pour $m = 1$. (On remarquera que pour $m = 1$, la condition d'involutivité est automatiquement réalisée.)

Supposons maintenant que $m > 1$ et raisonnons par récurrence. Supposons que la proposition 6 a été prouvée pour des sous-fibrés de fibres de dimension $m - 1$ et considérons un sous-fibré involutif E de fibres de dimension m.

Lemme 3. *Sur la variété M il existe une carte $(U, x^1, \ldots, x^n)$ avec $a_0 \in U$ et une base $X_1, \ldots, X_m$ du module $\mathfrak{a}(E|_U)$ telles que*

a) $X_m x^1 = \ldots = X_m x^{n-1} = 0, X_m x^n = 1$, *c'est-à-dire que* $X_m = \frac{\partial}{\partial x^n}$;

b) $X_1 x^n = \ldots = X_{m-1} x^n = 0$, *c'est-à-dire que les champs* $X_1, \ldots, X_{m-1}$ *ne s'expriment qu'en fonction des champs* $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^{n-1}}$;

c) *pour* $x^n = 0$, *les fonctions* $X_1 x^j, \ldots, X_{m-1} x^j$, $m \leqslant j \leqslant n$, *de* $x^1, \ldots, x^{n-1}$ *sont identiquement nulles.*

D é m o n s t r a t i o n. Supposons tout d'abord que $(U, x_1, \ldots, x_n)$ est une carte (avec $a_0 \in U$) pour laquelle le module $\mathfrak{a}(E|_U)$ est libre et soit $X_1, \ldots, X_m$ une base du module $\mathfrak{a}(E|_U)$. Le champ de vecteurs X_m engendre un sous-fibré pour lequel $m = 1$; en appliquant à ce fibré la partie de la proposition 7 qui a été prouvée, on trouve une carte (que l'on désignera encore par $(U, x^1, \ldots, x^n)$) vérifiant la condition a).

Pour satisfaire la condition b), on remplacera les champs $X_1, \ldots, X_{m-1}$ par les champs

$$X_1 = (X_1 x^n) X_m, \ldots, X_{m-1} - (X_{m-1} x^n) X_m.$$

Il est évident que ces champs (que l'on désignera encore par $X_1, \ldots, X_{m-1}$) forment par adjonction du champ X_m une base de $\mathfrak{a}(E|_U)$ et satisfont à la condition b).

La condition c) est plus difficile à réaliser.

Soit $W \subset \mathbb{R}^{n-1}$ l'ensemble de la première partie de la démonstration ($w \in W$ si et seulement si $(w, 0) \in h(U)$). Identifiant W à $W \times 0$, considérons la restriction $\varphi : W \to U$ à W du difféomorphisme h^{-1} réciproque du difféomorphisme de carte $h : U \to h(U) \subset \mathbb{R}^n$. Pour tout point $w \in W$, la différentielle $(d\varphi)_w$ de l'application φ est un monomorphisme de l'espace vectoriel $\mathsf{T}_w(W) = \mathbb{R}^{n-1}$ sur

le sous-espace (de l'espace vectoriel $\mathsf{T}_a(M)$, $a = \varphi(w)$) engendré par les vecteurs $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots, \left(\frac{\partial}{\partial x^{n-1}}\right)_a$. Donc, il existe sur W des champs de vecteurs

$$Y_1, \ldots, Y_{m-1}$$

définis de façon unique qui sont φ-liés aux champs $X_1, \ldots, X_{m-1}$ (pour lesquels nous avons supposé remplies les conditions a) et b)). Pour tout point $w \in W$, en notant $(\varphi^*E)_w$ l'enveloppe linéaire des vecteurs $(Y_1)_w, \ldots, (Y_{m-1})_w$, on obtient au-dessus de W un sous-fibré φ^*E du fibré tangent $\mathsf{T}(W)$ tel que le morphisme $\mathsf{T}(\varphi)\colon \mathsf{T}(W) \to \mathsf{T}(M)$ de fibrés vectoriels l'envoie dans le sous-fibré E.

Par construction les champs de vecteurs $Y_1, \ldots, Y_{m-1}$ forment une base du $\mathcal{F}(W)$-module $\mathfrak{a}(\varphi^*E)$. Ceci étant, puisque pour tout point $w \in W$

$$(d\varphi)_w [Y_i, Y_j]_w = [X_i, X_j]_{\varphi(w)} \in E_{\varphi(w)},$$

il vient $[Y_i, Y_j]_w \in (\varphi^*E)_w$, et, par suite, $[Y_i, Y_j] \in \mathfrak{a}(\varphi^*E)$. Ceci signifie que le sous-fibré φ^*E est involutif.

Donc, d'après l'hypothèse de récurrence, il existe sur W (après une éventuelle réduction de U, donc de W) des coordonnées (curvilignes) $w^1, \ldots, w^{n-1}$ telles que les champs de vecteurs $\frac{\partial}{\partial w^1}, \ldots$ $\ldots, \frac{\partial}{\partial w^{n-1}}$ engendrent le sous-module $\mathfrak{a}(\varphi^*F)$. Comme cette propriété caractérise aussi les champs $Y_1, \ldots, Y_{m-1}$, il s'ensuit de là que les champs $Y_1, \ldots, Y_{m-1}$ ne s'expriment qu'en fonction de $\frac{\partial}{\partial w^1}, \ldots, \frac{\partial}{\partial w^{m-1}}$, et, par suite, leurs composantes $Y_1 w^j, \ldots$ $\ldots, Y_{m-1} w^j$ sont nulles dans la base $\frac{\partial}{\partial w^1}, \ldots, \frac{\partial}{\partial w^{n-1}}$ pour $m \leqslant j \leqslant n$.

Sans nuire à la généralité on peut admettre $h(U) = W \times I$, où I est un intervalle de $\mathbb{R}$, et, par suite, en posant $h(u) = (w, x^n)$, on peut considérer que les coordonnées locales de tout point $u \in U$ sont les coordonnées $w^1, \ldots, w^{n-1}$ du point $w \in W$ et la coordonnée $x^n \in I$. En d'autres termes, nous introduisons dans U de nouvelles coordonnées locales $y^1 = w^1, \ldots, y^{n-1} = w^{n-1}$, $y^n = x^n$, en adoptant pour difféomorphisme de carte $U \to h(U)$ le difféomorphisme réciproque du difféomorphisme $(\varphi \circ k) \times \mathrm{id}$, où $k\colon W \to W'$ est un difféomorphisme déterminant les coordonnées locales $w^1, \ldots, w^{n-1}$ dans W. Les coordonnées $y^1, \ldots, y^n$ sont telles que $y^n = x^n$, et $y^1, \ldots, y^{n-1}$ ne dépendent pas de x^n, de sorte que

$$\frac{\partial y^1}{\partial x^n} = 0, \ \ldots, \ \frac{\partial y^{n-1}}{\partial x^n} = 0$$

et

$$\frac{\partial y^n}{\partial x^1} = 0, \ \ldots, \ \frac{\partial y^n}{\partial x^{n-1}} = 0, \quad \frac{\partial y^n}{\partial x^n} = 1.$$

Donc, les composantes $Xy^1, \ldots, Xy^n$ d'un champ de vecteurs X dans les coordonnées $y^1, \ldots, y^n$ s'expriment en fonction de ses composantes $X^1 = Xx^1, \ldots, X^n = Xx^n$ dans les coordonnées $x^1, \ldots$ $\ldots, x^n$ au moyen des formules

$$Xy^i = X^j \frac{\partial y^i}{\partial x^j} =$$

$$= \begin{cases} X^1 \dfrac{\partial y^i}{\partial x^1} + \ldots + X^{n-1} \dfrac{\partial y^i}{\partial x^{n-1}} & \text{pour } i = 1, \ldots, n-1, \\ X^n & \text{pour } i = n. \end{cases}$$

En particulier, pour le champ $X = X_m$ (pour lequel par hypothèse $X^1 = 0, \ldots, X^{n-1} = 0, X^n = 1$), on déduit de là que $X_m y^1 =$ $= 0, \ldots, X_m y^{n-1} = 0$, $X_m y^n = 1$, c'est-à-dire que $X_m = \dfrac{\partial}{\partial y^n}$. Pour les champs $X_1, \ldots, X_{m-1}$, il résulte de là que $X_1 y^n = \ldots$ $\ldots = X_{m-1} y^n = 0$, c'est-à-dire que ces champs ne s'expriment linéairement qu'en fonction des champs $\dfrac{\partial}{\partial y^1}, \ldots, \dfrac{\partial}{\partial y^{n-1}}$. Ceci signifie que les conditions a) et b) sont encore remplies dans la carte $(U, y^1, \ldots, y^n)$.

D'autre part, pour $y^n = 0$, c'est-à-dire en un point de la forme $u = \varphi(w)$, $w \in W$, les valeurs $(X_i)_u$ du champ de vecteurs X_i, $i = 1, \ldots, n-1$, sont justiciables de la formule

$$(X_i)_u = (d\varphi)_w Y_i,$$

qui exprime la φ-liaison des champs X_i et Y_i. Appliquée aux fonctions $X_i y^j$, cette formule montre que la valeur

$$(X_i y^j)(u) = (X_i)_u x^j$$

de la fonction $X_i y^j$ en $u \in U$ avec $y^n(u) = 0$ est égale à la valeur de la fonction $Y_i y^j$ en w. Comme $Y_i y^j = 0$ pour $m \leqslant j \leqslant n-1$, ceci prouve que $X_i y^j|_{y^n=0} = 0$ pour $m \leqslant j \leqslant n-1$. Donc, la condition c) (dans laquelle il faut remplacer $x^1, \ldots, x^n$ par $y^1, \ldots$ $\ldots, y^n$) est remplie pour la carte $(U, y^1, \ldots, y^n)$. □

Ce lemme entraîne sans peine la proposition 7.

Supposons maintenant que la carte $(U, x^1, \ldots, x^n)$ et les champs $X_1, \ldots, X_m$ vérifient les conditions du lemme 3.

Considérons le crochet de Lie $[X_m, X_i]$, $i = 1, \ldots, m-1$. En vertu de la condition d'involutivité, ce champ de vecteurs est situé dans $\mathfrak{a}(E|_U)$, et, par suite, il existe sur U des fonctions différentiables $c_i^1, \ldots, c_i^m$ telles que

$$[X_m, X_i] = c_i^1 X_1 + \ldots + c_i^m X_m, \qquad i = 1, \ldots, m-1,$$

donc

$$[X_m, X_i]\, x^j = c_i^1 X_1 x^j + \ldots + c_i^m X_m x^j$$

pour tout $j = 1, \ldots, n$. Ceci étant, si $j \neq n$ (le seul cas intéressant), alors $X_m x^j = 0$. On peut donc admettre que le dernier terme du second membre est de la forme $c_i^{m-1} X_{m-1} x^j$. Par ailleurs, par définition du crochet de Lie

$$[X_m, X_i] x^j = X_m (X_i x^j) - X_i (X_m x^j) = \frac{\partial X_i x^j}{\partial x^m},$$

puisque par hypothèse $X_m = \frac{\partial}{\partial x^m}$. Ceci signifie que pour tout $j = 1, \ldots, n-1$ (de même, évidemment, que pour $j = n$), les fonctions

$$z_1 = X_1 x^j, \ldots, z_{m-1} = X_{m-1} x^j,$$

regardées comme des fonctions de $t = x^n$, sont solutions du système d'équations différentielles

$$\frac{dz_i}{dt} = c_i^1 z_1 + \ldots + c_i^{m-1} z_{m-1}, \quad i = 1, \ldots, m-1.$$

Or, d'après la condition c) du lemme 3, les fonctions $z_1, \ldots, z_{m-1}$ sont nulles pour $t = 0$. Elles le sont donc pour tout t.

Ceci prouve que sur U les champs $X_1, \ldots, X_{m-1}$ s'expriment en fonction du champ $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^{m-1}}$ et, par suite, les champs $X_1, \ldots, X_m$, en fonction du champ $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^{m-1}}, \frac{\partial}{\partial x^n}$. Donc, les derniers champs forment une base du module $c(E|_U)$.

Pour achever la récurrence, il reste à changer les noms des coordonnées x^n et x^m.

Ceci prouve la proposition 6 et partant la proposition 5 et le lemme 2. □

Corollaire. *Un sous-fibré d'un fibré tangent est intégrable si et seulement s'il est complètement intégrable.* □

LEÇON 8

Revêtements.— Sections des revêtements.— Revêtements pointés.— Coamalgames.— Espaces simplement connexes.— Morphismes de revêtements.— Relation de préordre dans la catégorie des revêtements pointés.— Existence de revêtements simplement connexes.— Problèmes de justification.— Fonctorialité d'un revêtement universel.

Passons maintenant à l'étude du foncteur de localisation GR_0-DIFF→ → GR-LOC. L'étude de ce foncteur est basée sur d'autres idées et méthodes liées essentiellement aux espaces de revêtement. L'édification habituelle de la théorie des espaces de revêtement à l'aide de la notion de chemins homotopes a fait l'objet de nombreux manuels et de monographies tant dans la perspective de son application à la théorie des groupes de Lie (cf. par exemple [8], chap. 9) qu'indépendamment d'elle (cf. par exemple [13], chap. 5). On se propose d'exposer une construction plus élémentaire de cette théorie que l'on doit à Chevalley, construction qui s'appuiera uniquement sur des considérations topologiques simples. On appliquera ensuite les résultats acquis à l'étude du foncteur de localisation.

Définition 1. Soit $\pi: \widetilde{X} \to X$ une application surjective continue d'un espace topologique $\widetilde{X}$ sur un espace topologique X, et soit $U \subset X$ un ouvert de X. On dit que U est *revêtu sans pli* par l'application π si l'image réciproque $\pi^{-1}(U)$ est la réunion d'ensembles ouverts disjoints homéomorphes chacun à U par l'application π. Une application π est un *revêtement* de l'espace X si les espaces X et $\widetilde{X}$ sont connexes et si tout point de X possède un voisinage revêtu sans pli par l'application π. L'espace $\widetilde{X}$ s'appelle alors *espace de revêtement.*

On a parfois affaire à des applications surjectives $\pi: \widetilde{X} \to X$ pour lesquelles l'espace X est connexe et chacun de ses points possède

un voisinage revêtu sans pli par π, mais l'espace $\tilde{X}$ n'est généralement pas connexe. De telles applications seront dites *revêtements faibles*.

Si U est revêtu sans pli par une application π, il en sera de même de tout ouvert $V \subset U$. Donc, *pour tout revêtement* $\pi: \tilde{X} \to X$ *(généralement faible), les ouverts revêtus sans pli* $U \subset X$ *forment une base dans l'espace* X.

Puisque tout point $\tilde{x} \in \tilde{X}$ appartient à l'image réciproque d'un point $x \in X$, les ouverts $\tilde{U} \subset \tilde{X}$, qui sont homéomorphes par l'application π à des ouverts $U \subset X$, forment une base dans l'espace $\tilde{X}$, autrement dit tout ouvert de $\tilde{X}$ est la réunion de tels ensembles.

Une application continue $\pi: \tilde{X} \to X$ admettant la dernière propriété s'appelle *homéomorphisme local*. Donc, *tout revêtement (faible) est un homéomorphisme local.*

Signalons que la réciproque est fausse. Un exemple d'homéomorphisme local qui n'est pas un revêtement nous est donné par la restriction d'un revêtement $\pi: \tilde{X} \to X$ au sous-espace $\tilde{X} \setminus \{\tilde{x}_0\}$, où $\tilde{x}_0 \in \tilde{X}$ est un point arbitraire (même si cette restriction est surjective et le sous-espace $\tilde{X} \setminus \{\tilde{x}_0\}$, connexe).

Tout homéomorphisme local est visiblement une application ouverte (c'est-à-dire envoie ouvert dans ouvert). Donc, *tout revêtement (faible) est une application ouverte.*

D'une façon générale, l'image réciproque d'un ensemble $U \subset X$ revêtu sans pli ne se représente pas de façon unique par la réunion d'ouverts $\tilde{U}_\alpha$ disjoints homéomorphes chacun à U. Mais *cette représentation est unique si U est connexe*, puisque dans ce cas les ensembles $\tilde{U}_\alpha$ peuvent être traités comme les composantes connexes de l'ensemble $\pi^{-1}(U)$ (étant homéomorphe à U, chacun de ces ensembles est connexe et de plus ils sont ouverts et disjoints).

De nombreuses propriétés topologiques importantes de l'espace X sont héritées par tout espace de revêtement $\tilde{X}$. Ainsi, du fait que le revêtement $\pi: \tilde{X} \to X$ est un homéomorphisme local, il s'ensuit immédiatement que *l'espace $\tilde{X}$ est localement connexe s'il en est de même de l'espace X.*

De façon analogue, il est aisé de voir que *si un espace X est séparé, il en est de même de chacun de ses espaces de revêtement $\tilde{X}$.* En effet, soient $\tilde{x}_1$, $\tilde{x}_2$ deux points distincts de $\tilde{X}$. Si $\pi(\tilde{x}_1) = \pi(\tilde{x}_2)$, les points $\tilde{x}_1$ et $\tilde{x}_2$ appartiennent par hypothèse à des ouverts disjoints.

Si $\pi(\tilde{x}_1) \neq \pi(\tilde{x}_2)$, les points $\pi(\tilde{x}_1)$ et $\pi(\tilde{x}_2)$ possèdent des voisinages U_1 et U_2 disjoints, puisque l'espace X est séparé. Les images réciproques $\pi^{-1}(U_1)$ et $\pi^{-1}(U_2)$ de ces voisinages seront des ensembles ouverts disjoints qui contiennent $\tilde{x}_1$ et $\tilde{x}_2$. □

Soit $U \subset X$ un ouvert revêtu sans pli par une application $\pi: \tilde{X} \to X$. Considérons des ouverts disjoints $\tilde{U}_\alpha \subset \tilde{X}$ sur lesquels π est un homéomorphisme et dont la réunion est l'ensemble $\pi^{-1}(U)$. Par définition, pour tout α l'homéomorphisme $\sigma_\alpha: U \to \tilde{U}_\alpha$ réciproque de l'homéomorphisme $\pi|_{\tilde{U}_\alpha}: \tilde{U}_\alpha \to U$ sera une section de l'application π au-dessus de U (plus exactement, c'est sa composée avec l'injection $\tilde{U}_\alpha \to \tilde{X}$ qui sera une section).

Par ailleurs, on sait que si U est connexe, les ensembles U_α sont définis de façon unique; de plus, pour se donner l'un d'eux, il suffit d'en indiquer un point seulement (puisque ces ensembles sont les composantes de l'ensemble $\pi^{-1}(U)$). Donc, les sections σ_α sont définies aussi de façon unique. Cela signifie que si dans U nous choisissons un point x_0, alors pour tout point $\tilde{x}_\alpha \in \pi^{-1}(x_0)$, il existera au-dessus de U une seule section $\sigma_\alpha: U \to \tilde{X}$ de l'application π pour laquelle $\sigma(\tilde{x}_\alpha) = x_0$.

Réciproquement, soit U un sous-ensemble connexe de X ayant la propriété suivante: pour tout point $x_0 \in U$ et tout point $\tilde{x}_\alpha \in \pi^{-1}(x_0)$, il existe une seule section $\sigma_\alpha: U \to \tilde{X}$ de l'application π au-dessus de U pour laquelle $\sigma_\alpha(x_0) = \tilde{x}_\alpha$. Soit un point arbitraire $\tilde{x} \in \pi^{-1}(U)$. Supposons que $x = \pi(\tilde{x})$. Par hypothèse, il existe une seule section $\sigma: U \to \tilde{X}$ de l'application π au-dessus de U telle que $\sigma(x) = \tilde{x}$. Soit $\sigma(x_0) = \tilde{x}_\alpha$. L'unicité de la section σ_α entraîne l'égalité $\sigma = \sigma_\alpha$ qui montre, en particulier, que $\tilde{x} \in \sigma_\alpha(U)$. Donc, l'ensemble $\pi^{-1}(U)$ est la réunion des ensembles $\sigma_\alpha(U)$. Si $\tilde{x} \in \sigma_{\alpha_1}(U) \cap \sigma_{\alpha_2}(U)$, alors $\tilde{x}_{\alpha_1} = \sigma_{\alpha_1}(x_0) = \sigma(x_0) = \sigma_{\alpha_2}(x_0) = \tilde{x}_{\alpha_2}$ et, par suite, $\sigma_{\alpha_1} = \sigma_{\alpha_2}$. Donc, les ensembles $\sigma_\alpha(U)$ sont disjoints. Comme ces ensembles sont connexes (puisque homéomorphes à l'ensemble connexe U), ceci prouve qu'ils sont les composantes de l'ensemble $\pi^{-1}(U)$. Si l'espace $\tilde{X}$ est localement connexe, ces composantes doivent être ouvertes. Donc, sous les conditions posées, l'ensemble $\pi^{-1}(U)$ est la réunion d'ouverts disjoints homéomorphes chacun à U. Autrement dit, l'ensemble U est revêtu sans pli par l'application π.

Nous avons ainsi établi la proposition suivante:

Proposition 1. *Si l'espace $\widetilde{X}$ est localement connexe, un ouvert connexe $U \subset X$ est revêtu sans pli par l'application $\pi: \widetilde{X} \to X$ si et seulement si, pour tout point $x_0 \in U$ et tout point $\widetilde{x}_\alpha \in \pi^{-1}(x_0)$, il existe une seule section σ_α de l'application π au-dessus de U telle que $\sigma_\alpha(x_0) = \widetilde{x}_\alpha$.* □

Corollaire. *Soient U et V des ouverts connexes de l'espace X revêtus sans pli par l'application $\pi: \widetilde{X} \to X$. Si:*

a) *l'espace $\widetilde{X}$ est localement connexe;*

b) *l'intersection $U \cap V$ est connexe, alors l'ensemble $U \cup V$ est aussi revêtu sans pli par l'application π.*

Démonstration. Soient $x_0 \in U \cup V$ et $\widetilde{x}_\alpha \in \pi^{-1}(x_0)$. Il suffit de montrer qu'il existe une section $\sigma_\alpha: U \cup V \to \widetilde{X}$ de l'application π au-dessus de $U \cup V$, telle que $\sigma_\alpha(x_0) = \widetilde{x}_\alpha$ et que cette section est unique. Sous ces conditions, on peut, sans nuire à la généralité, admettre visiblement que $x_0 \in U$.

Par hypothèse, il existe au-dessus de U une section $\sigma'_\alpha: U \to \widetilde{X}$ de l'application π, telle que $\sigma'_\alpha(x_0) = \widetilde{x}_\alpha$ et cette section est unique. Choisissons un point quelconque $y_0 \in U \cap V$ et considérons le point $\widetilde{y}_0 = \sigma'_\alpha(y_0) \in \widetilde{X}$. Comme $y_0 \in V$, il existe au-dessus de V une seule section $\sigma''_\alpha: V \to \widetilde{X}$ de l'application π, telle que $\sigma''_\alpha(y_0) = \widetilde{y}_0$. Les restrictions à $U \cap V$ des sections σ'_α et σ''_α sont des sections au-dessus de $U \cap V$ qui envoient y_0 dans $\widetilde{y}_0$. Comme $U \cap V$ est connexe, il s'ensuit de là que $\sigma'_\alpha = \sigma''_\alpha$ sur $U \cap V$. Donc, les sections σ'_α et σ''_α définissent sur $U \cup V$ une application continue $\sigma_\alpha: U \cup \cup V \to \widetilde{X}$ qui est visiblement une section telle que $\sigma_\alpha(x_0) = \widetilde{x}_\alpha$. Ceci établit l'existence de la section σ_α. L'unicité est maintenant triviale. □

La proposition 1 et le corollaire sont valables notamment pour tout revêtement $\pi: \widetilde{X} \to X$ d'un espace X localement connexe (et connexe).

Dans les hypothèses de la proposition 1, l'image $\sigma_\alpha(U)$ de U par chaque application σ_α est un sous-ensemble ouvert de l'espace $\widetilde{X}$. La proposition suivante montre que ce fait revêt un caractère assez général:

Proposition 2. *Si l'espace X est localement connexe, pour tout revêtement $\pi: \widetilde{X} \to X$ chacune de ses sections $\sigma: V \to \widetilde{X}$ au-dessus d'un ouvert $V \subset X$ est une application ouverte.*

En particulier, l'ensemble $\sigma(V)$ *est ouvert dans* $\tilde{X}$, *et* σ *est un homéomorphisme de* V *sur* $\sigma(V)$ (*admettant l'homéomorphisme réciproque* $\pi|_{\sigma(V)}$).

D é m o n s t r a t i o n. Puisque l'application σ est une section au-dessus de chaque ouvert contenu dans V, il suffit de montrer que l'ensemble $\sigma(V)$ est ouvert dans $\tilde{X}$.

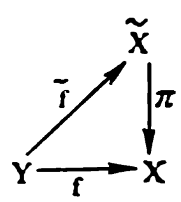
Soient $\tilde{x} \in \sigma(V)$ et $x \in V$, un point tel que $\sigma(x) = \tilde{x}$. Soient par ailleurs $U \subset V$ un voisinage connexe de x revêtu sans pli par l'application π, $\tilde{U}$ la composante de son image réciproque $\pi^{-1}(U)$ qui contient le point $\tilde{x}$. L'image $\sigma(U)$ du voisinage U par l'application σ est un ensemble connexe contenant le point $\tilde{x}$. Donc, $\sigma(U) \subset \subset \tilde{U}$, et, par suite, est définie l'application $(\pi|_{\tilde{U}}) \circ \sigma = \mathrm{id}$. Or, par hypothèse, l'application $\pi|_{\tilde{U}}: \tilde{U} \to U$ est un homéomorphisme. Donc, si $\sigma': U \to \tilde{U}$ est l'homéomorphisme réciproque, il vient $\sigma' \circ (\pi|_{\tilde{U}}) = \mathrm{id}$, et, par suite,

$$(\sigma' = \sigma' \circ (\pi|_{\tilde{U}} \circ \sigma) = (\sigma' \circ \pi|_{\tilde{U}}) \circ \sigma = \sigma.$$

En particulier, $\sigma(U) = \sigma'(U) = \tilde{U}$. Ceci signifie que le voisinage $\tilde{U}$ du point $\tilde{x} \in \sigma(V)$ est entièrement contenu dans $\sigma(V)$, de sorte que $\tilde{x}$ est un point intérieur de l'ensemble $\sigma(V)$. Donc, l'ensemble $\sigma(V)$ est ouvert. □

L'unicité des sections σ_α impliquée dans la proposition 1 peut effectivement être prouvée (notamment dans le cas seul intéressant pour nous où π est un revêtement d'un espace séparé). Du reste, il nous sera plus commode d'établir cette unicité dans une situation plus générale pour la description de laquelle nous introduisons la définition suivante:

Définition 2. Soient $f: Y \to X$ et $\pi: \tilde{X} \to X$ deux morphismes d'une catégorie $\mathbf{C}$. On dit qu'un morphisme $\tilde{f}: Y \to \tilde{X}$ de la catégorie $\mathbf{C}$ est un *relèvement* du morphisme f (relativement au morphisme π) si $f = \pi \circ \tilde{f}$, c'est-à-dire si le diagramme



est commutatif.

Les sections ne sont autres que les relèvements du morphisme

identique $X \to X$, et les sections au-dessus de $U \subset X$ (dans le cas où la catégorie C est la catégorie TOP des espaces topologiques), les relèvements de l'injection $U \to X$.

Proposition 3. *Soient $\pi: \tilde{X} \to X$ un revêtement de l'espace séparé X, $f: Y \to X$ une application continue d'un espace connexe Y dans l'espace X. Si deux relèvements $\tilde{f}, \tilde{f}': Y \to \tilde{X}$ sont confondus en un point au moins $y_0 \in Y$, ils le sont partout.*

Démonstration. Soit Y' l'ensemble des points $y \in Y$ tels que $\tilde{f}y = \tilde{f}'y$. L'ensemble Y' n'est pas vide (il contient le point y_0) et est fermé (car l'espace X est séparé). L'espace Y étant connexe par hypothèse, pour prouver la proposition 3, il suffit de montrer que l'ensemble Y' est également ouvert.

Soit $y \in Y'$, et soit U un voisinage du point $f(y) \in X$, revêtu sans pli par l'application π. Sous ces conditions, le point $\tilde{f}(y) = \tilde{f}'(y)$ possède dans $\tilde{X}$ un voisinage $\tilde{U}$ sur lequel π est un homéomorphisme $\tilde{U} \to U$. Les applications $\tilde{f}$ et $\tilde{f}'$ étant continues, le point y possède dans Y un voisinage V tel que $\tilde{f}(V) \subset \tilde{U}$ et $\tilde{f}'(V) \subset \tilde{U}$. Comme l'application π est un homéomorphisme sur $\tilde{U}$ et que $\pi \circ \tilde{f} = \pi \circ \tilde{f}'$, on a $\tilde{f} = \tilde{f}'$ sur V, c'est-à-dire que $V \subset Y'$. L'ensemble Y' est donc ouvert. □

Nous aurons besoin de la définition topologique suivante :

Définition 3. Un espace topologique X est *pointé* si on y a distingué un point base x_0. On appelle *application* $(X, x_0) \to (Y, y_0)$ d'un espace pointé (X, x_0) dans un espace pointé (Y, y_0) une application continue $X \to Y$ envoyant x_0 dans y_0.

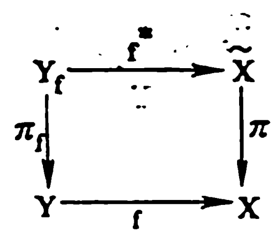
Il est évident que les espaces pointés et leurs applications forment une catégorie que nous noterons TOP$^\bullet$.

On appelle *revêtement pointé* une application $\pi: (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ de la catégorie TOP$^\bullet$ qui est un revêtement en tant qu'application $\tilde{X} \to X$ de la catégorie TOP.

Traduite dans ce langage, la proposition 3 affirme que *toute application $f: (Y, y_0) \to (X, x_0)$ d'un espace pointé connexe (Y, y_0) dans un espace pointé séparé (X, x_0) admet au plus un relèvement $\tilde{f}: (Y, y_0) \to (\tilde{X}, \tilde{x}_0)$ relativement au revêtement pointé $\pi: (\tilde{X}, \tilde{x}_0) \to (X, x_0)$.*

Dans la suite, pour alléger les notations, pour (X, x_0), $(\tilde{X}, \tilde{x}_0)$, etc., on écrira simplement X, $\tilde{X}$, etc., en n'indiquant explicitement les points bases que dans le cas où une confusion serait à craindre.

Soient encore $f: Y \to X$ et $\pi: \tilde{X} \to X$ des morphismes d'une catégorie **C** et supposons que ces morphismes sont inclus dans un diagramme commutatif de la forme



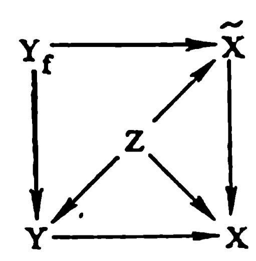
Ce diagramme s'appelle *carré universel* et le morphisme $\pi_f: Y_f \to Y$ (ou l'objet Y_f), *coamalgame* des morphismes f et π (ou des objets Y et $\tilde{X}$ au-dessus de l'objet X) si pour tout objet Z et tous morphismes $g_1: Z \to Y$ et $g_2: Z \to \tilde{X}$ vérifiant la relation

(1) $$f \circ g_1 = \pi \circ g_2,$$

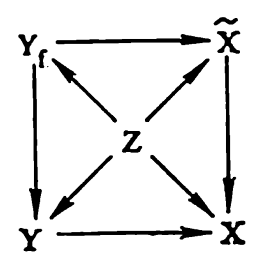
il existe un seul morphisme $g: Z \to Y_f$, pour lequel

(2) $$\pi_f \circ g = g_1, \qquad f^* \circ g = g_2,$$

ou, en d'autres termes, si tout diagramme commutatif de la forme

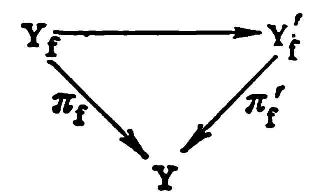


se complète de façon unique à un diagramme commutatif de la forme



Deux coamalgames quelconques Y_f et Y'_f des morphismes f et π sont canoniquement isomorphes: il existe un isomorphisme $Y_f \to$

$\to Y'_f$ et un seul rendant commutatif le diagramme



La propriété fondamentale du coamalgame $\pi_f: Y_f \to Y$ est que *ses sections $g: Y \to Y_f$ se trouvent de façon naturelle en correspondance biunivoque avec les relèvements $\tilde{f}: Y \to \tilde{X}$ du morphisme f.* En effet, toute section g détermine un relèvement $f^* \circ g$ et pour chaque relèvement $\tilde{f}$, les morphismes $g_1 = \text{id}$ et $g_2 = \tilde{f}$ vérifient (pour $Z = Y$) les conditions (1), et, par suite, définissent un morphisme $g: Y \to Y_f$ qui est (en vertu de la première relation (2)) une section du morphisme π_f. □

De ce point de vue les relèvements se réduisent à leur cas particulier: les sections.

Cette réduction suppose évidemment l'existence du coamalgame π_f. Il s'avère que dans le cas de la catégorie TOP (ou TOP$^\cdot$) qui nous intéresse, *le coamalgame $\pi_f: Y_f \to Y$ existe pour toutes applications continues $f: Y \to X$ et $\pi: \tilde{X} \to X$.* L'espace Y_f correspondant est dans ce cas un sous-espace, du produit direct $Y \times \tilde{X}$, des points $(y, \tilde{x})$ tels que $f(y) = \pi(\tilde{x})$, et les applications π_f et f^* sont les restrictions à Y_f des projections de ce produit sur ses facteurs. Que cela nous donne bien un coamalgame se vérifie directement: en effet, si $f \circ g_1 = \pi \circ g_2$, l'application $g = g_1 \times g_2: Z \to Y \times \tilde{X}$ est une application dans Y_f et satisfait les relations $\pi_f \circ g = g_1$ et $f^* \circ g = g_2$; par ailleurs, π_f et f^* étant des projections, ces relations déterminent une seule application g. □

Lemme 1. *Si une application $\pi: \tilde{X} \to X$ est un revêtement (faible), pour toute application $f: Y \to X$, le coamalgame*

$$\pi_f: Y_f \to Y$$

sera un revêtement faible.

D é m o n s t r a t i o n. L'application π_f est surjective, puisque pour tout point $y \in Y$ il existe un point $\tilde{x} \in \tilde{X}$ tel que $\pi(\tilde{x}) = f(y)$. Soit U un sous-ensemble ouvert de l'espace X, revêtu sans pli par l'application π, et soit $V = f^{-1}(U)$ l'image réciproque de U par f. Il est clair que

$$\pi_f^{-1}(V) = Y_f \cap (V \times \pi^{-1}(U)).$$

Donc, si

$$\pi^{-1}(U) = \bigcup_\alpha \widetilde{U}_\alpha,$$

où $\widetilde{U}_\alpha$ sont des sous-ensembles ouverts disjoints de l'espace $\widetilde{X}$, homéomorphes à U, alors

$$\pi_f^{-1}(V) = \bigcup_\alpha V_\alpha,$$

où

$$V_\alpha = Y_f \cap (V \times \widetilde{U}_\alpha).$$

Les ensembles V_α sont ouverts, disjoints et homéomorphes à V par l'application π_f. Donc, l'ensemble V est revêtu sans pli par l'application π_f. Pour achever la démonstration, il reste à remarquer que les ensembles V revêtent Y tout entier. □

Si donc nous voulons obtenir un revêtement, il nous suffit de passer de l'espace Y_f à l'une de ses composantes. Le lemme suivant nous montre qu'on obtient bien un revêtement:

Lemme 2. *Si l'espace X est connexe et localement connexe, alors pour tout revêtement faible $\pi: \widetilde{X} \to X$ et toute composante $\widetilde{X}_0$ de l'espace $\widetilde{X}$, l'application*

$$\pi_0 = \pi|_{\widetilde{X}_0}: \widetilde{X}_0 \to X$$

est un revêtement.

D é m o n s t r a t i o n. Soit U un sous-ensemble ouvert connexe de l'espace X, revêtu sans pli par l'application π. Les composantes de l'ensemble $\pi^{-1}(U)$ qui coupent $\widetilde{X}_0$ sont nécessairement situées dans $\widetilde{X}_0$ (en vertu de la connexité). Donc, si $\pi_f(\widetilde{X}_0) \cap U \neq \varnothing$, alors $U \subset \pi(\widetilde{X}_0)$. Les ensembles U formant une base de l'espace X, ceci montre que l'ensemble $\pi(\widetilde{X}_0)$ est ouvert et fermé. Donc, il revêt X tout entier, de sorte que l'application π_0 est surjective. Comme $\pi_0^{-1}(U)$ est la réunion des composantes de l'ensemble $\pi^{-1}(U)$ situées dans $\widetilde{X}_0$, l'application π_0 revêt sans pli l'ensemble U. Donc, π_0 est un revêtement. □

Voyons maintenant le problème de l'existence des relèvements (et des sections). Il nous faut commencer de loin.

Il est clair que tout homéomorphisme $\pi: \widetilde{X} \to X$ est un revêtement.

Définition 4. Un revêtement $\pi: \widetilde{X} \to X$ est *trivial* s'il est un homéomorphisme de $\widetilde{X}$ sur X.

Proposition 4. *Un revêtement* $\pi: \tilde{X} \to X$ *d'un espace localement connexe* (*et connexe*) $\tilde{X}$ *est trivial si et seulement s'il possède une section* $\sigma: X \to \tilde{X}$ *au-dessus de l'espace* X *tout entier.*

D é m o n s t r a t i o n. Si le revêtement π est trivial, l'homéomorphisme réciproque $X \to \tilde{X}$ sera la section σ (que X soit localement connexe ou non).

Réciproquement, supposons que le revêtement $\pi: \tilde{X} \to X$ possède une section $\sigma: X \to \tilde{X}$. En vertu de la proposition 2, cette section est un homéomorphisme de l'espace X sur un ouvert $\sigma(X) \subset \tilde{X}$, admettant pour réciproque $\pi|_{\sigma(X)}: \sigma(X) \to X$. Pour prouver la proposition 4, il suffit de montrer que $\sigma(X) = \tilde{X}$. L'espace $\tilde{X}$ étant connexe, et l'ensemble $\sigma(X)$ ouvert et non vide, ceci sera prouvé si l'on montre que l'ensemble $\sigma(X)$ est fermé dans $\tilde{X}$.

Soit $\tilde{x}$ un point quelconque de l'adhérence $\overline{\sigma(X)}$ de l'ensemble $\sigma(X)$, et soit U un voisinage du point $x = \pi(\tilde{x})$ revêtu sans pli par l'application π. Considérons un ensemble ouvert $\tilde{U}$ contenant le point $\tilde{x}$ et homéomorphe au voisinage U. Comme $\tilde{x} \in \overline{\sigma(X)}$, l'intersection $\tilde{U} \cap \sigma(X)$ n'est pas vide. Soit $\tilde{y} \in \tilde{U} \cap \sigma(X)$. L'application $\pi|_{\sigma(X)}: \sigma(X) \to X$ étant un homéomorphisme, il existe dans $\sigma(X)$ un ensemble ouvert $\tilde{U}'$ homéomorphe à U et contenant le point $\tilde{y}$. Comme U est parfaitement revêtu, les ensembles $\tilde{U}$ et $\tilde{U}'$ soit sont confondus, soit sont disjoints. Or, ils ont un point commun $\tilde{y}$. Donc $\tilde{U}' = \tilde{U}$, et, par suite, $\tilde{x} \in \tilde{U} = \tilde{U}' \subset \sigma(X)$. Donc, l'ensemble $\sigma(X)$ est fermé. □

Définition 5. Un espace connexe X est *simplement connexe* si n'importe lequel de ses revêtements est trivial.

Le théorème suivant souligne l'importance des espaces simplement connexes pour le problème d'existence des relèvements.

Théorème 1. *Soit* $\pi: (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ *un revêtement pointé d'un espace séparé pointé* (X, x_0), *et soit* (Y, y_0) *un espace pointé connexe, localement et simplement connexe. Alors, pour toute application* $f: (Y, y_0) \to (X, x_0)$ *il existe un relèvement unique* $\tilde{f}: (Y, y_0) \to (\tilde{X}, \tilde{x}_0)$.

D é m o n s t r a t i o n. Par définition $(y_0, \tilde{x}_0) \in Y_f$. Soit $(Y_f)_0$ la composante de l'espace Y_f qui contient le point $(y_0, \tilde{x}_0)$. D'après

le lemme 2, l'application $(\pi f)_0 = \pi_f|_{(Y_f)_0} : (Y_f)_0 \to Y$ est un revêtement. Donc, en vertu de la simple connexité de l'espace Y, cette application est un homéomorphisme. Si $(\pi_f)_0^{-1}: Y \to (Y_f)_0$ est l'homéomorphisme réciproque, l'application

$$\tilde{f} = f^* \circ (\pi_f)_0^{-1}: Y \to \tilde{X}$$

sera un relèvement de l'application f tel que $\tilde{f}(y_0) = \tilde{x}_0$. L'unicité du relèvement $\tilde{f}$ est affirmée par la proposition 3. □

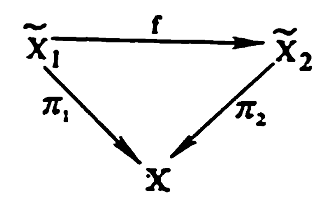
Corollaire 1. *Tout revêtement $\pi: \tilde{X} \to X$ d'un espace séparé connexe et localement connexe X revêt sans pli chaque sous-ensemble ouvert simplement connexe $U \subset X$.*

D é m o n s t r a t i o n. Il suffit d'utiliser la proposition 1. □

Corollaire 2. *Un espace X séparé connexe et localement connexe est simplement connexe s'il est la réunion de deux ouverts U et V connexes et simplement connexes dont l'intersection $U \cap V$ est connexe.*

D é m o n s t r a t i o n. D'après le corollaire 1, tout revêtement $\pi: \tilde{X} \to X$ revêt sans pli aussi bien U que V. Donc (cf. corollaire 1 de la proposition 1), il revêt sans pli $X = U \cup V$ et, par suite, est trivial. □

Définition 6. On appelle *morphisme* d'un revêtement $\pi_1: \tilde{X}_1 \to X$ dans un revêtement $\pi_2: \tilde{X}_2 \to X$ une application continue $f: \tilde{X}_1 \to \tilde{X}_2$ telle que l'on a le diagramme commutatif



Il est évident que les revêtements (d'un espace connexe X donné) et leurs morphismes forment une catégorie que l'on désignera par COV(X).

On définit de façon analogue la catégorie des *revêtements pointés* COV(X, x_0) dont les morphismes sont les morphismes de la catégorie COV(X) qui sont simultanément des applications d'espaces pointés.

On désignera la catégorie COV(X, x_0) aussi par COV$^\bullet$(X).

Il est clair qu'*un morphisme $f: \tilde{X}_1 \to \tilde{X}_2$ de la catégorie* COV(X) *(ou* COV(X, x_0)*) est un isomorphisme (possède un morphisme réciproque) si et seulement s'il est un homéomorphisme de l'espace $\tilde{X}_1$ sur l'espace $\tilde{X}_2$.*

10—01642

Un revêtement $\pi: \tilde{X} \to X$ est trivial (au sens de la définition 4) si et seulement s'il est isomorphe dans la catégorie COV(X) au revêtement identique id: $X \to X$.

Lemme 3. *Si un espace connexe X est localement connexe, tout morphisme $f: \tilde{X}_1 \to \tilde{X}_2$ de la catégorie* COV(X) (*ou de la catégorie* COV(X, x_0)) *est un revêtement* (*de l'espace* $\tilde{X}_2$).

D é m o n s t r a t i o n. Soit $\{U_\alpha\}$ une base de l'espace X formée d'ensembles ouverts connexes revêtus sans pli simultanément par des applications $\pi_1: \tilde{X}_1 \to X$ et $\pi_2: \tilde{X}_2 \to X$, et soient $\tilde{U}^{(1)}_{\alpha,\beta}$ les composantes de l'image réciproque $\pi_1^{-1}(U_\alpha)$ d'un ensemble U_α par l'application π_1, $\tilde{U}^{(2)}_{\alpha,\gamma}$ les composantes de l'image réciproque $\pi_2^{-1}(U_\alpha)$ de l'ensemble U_α par l'application π_2. Comme $\pi_1 = \pi_2 \circ f$, chaque ensemble $\tilde{U}^{(1)}_{\alpha,\beta}$ est homéomorphe par f à un ensemble $\tilde{U}^{(2)}_{\alpha,\gamma}$. Donc, si pour certains α et γ, l'ensemble $\tilde{U}^{(2)}_{\alpha,\gamma}$ coupe le sous-espace $f(\tilde{X}_1)$, il y sera nécessairement contenu: $\tilde{U}^{(2)}_{\alpha,\gamma} \subset f(\tilde{X}_1)$. Comme les ensembles $\tilde{U}^{(2)}_{\alpha,\gamma}$ forment une base dans l'espace $\tilde{X}_2$, ceci n'est possible que si le sous-espace $f(\tilde{X}_1)$ est ouvert et fermé. Donc, $f(\tilde{X}_1) = \tilde{X}_2$, et, par suite, l'application f est une surjection.

On voit par ailleurs que l'image réciproque $f^{-1}(\tilde{U}^{(2)}_{\alpha,\gamma})$ de tout ensemble $\tilde{U}^{(2)}_{\alpha,\gamma}$ par l'application f est la réunion de certains (en fait, de tous les) ensembles $\tilde{U}^{(1)}_{\alpha,\gamma}$ et, de plus, l'application f est un homéomorphisme sur chacun d'eux. Donc, tout ensemble $\tilde{U}^{(2)}_{\alpha,\gamma}$ est revêtu sans pli par l'application f. Donc, f est un revêtement. □

Chaque morphisme f d'un revêtement $\pi_1: \tilde{X}_1 \to X$ dans un revêtement $\pi_2: \tilde{X}_2 \to X$ n'est autre qu'un relèvement de l'application $\pi_1: \tilde{X}_1 \to X$ par rapport à l'application $\pi_2: \tilde{X}_2 \to X$. Donc, d'après la proposition 3, *si l'espace X est séparé, pour deux revêtements pointés quelconques $\pi_1: (\tilde{X}, \tilde{x}_1) \to (X, x_0)$ et $\pi_2: (\tilde{X}_2, \tilde{x}_2) \to (X, x_0)$, il existe dans la catégorie* COV(X, x_0) *au plus un morphisme du revêtement π_1 dans le revêtement π_2.*

En particulier, pour tout revêtement $\pi: (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ il n'existe qu'un seul morphisme identique $\pi \to \pi$.

On écrira $\pi_1 \geqslant \pi_2$ si le morphisme $\pi_1 \to \pi_2$ existe. Il est évident que la relation ainsi définie sur l'ensemble des revêtements pointés

de l'espace (X, x_0) est réflexive et transitive, c'est-à-dire est une *relation de préordre*.

On dira que des revêtements π_1 et π_2 sont *équivalents* si l'on a simultanément $\pi_1 \geqslant \pi_2$ et $\pi_2 \geqslant \pi_1$. Il est évident que la relation $\geqslant$ induit une *relation d'ordre* sur les classes des revêtements équivalents. Ceci étant, il est immédiat de voir que *si l'espace X est séparé, les revêtements π_1 et π_2 sont équivalents si et seulement s'ils sont isomorphes*. En effet, les revêtements isomorphes sont visiblement équivalents. Réciproquement, si des revêtements π_1 et π_2 sont équivalents, c'est-à-dire s'il existe des morphismes $f: \pi_1 \to \pi_2$ et $g: \pi_2 \to \pi_1$, alors, d'après l'unicité, les morphismes $f \circ g: \pi_2 \to \pi_2$ et $g \circ f: \pi_1 \to \pi_1$ seront des morphismes identiques $\pi_2 \to \pi_2$ et $\pi_1 \to \pi_1$, et, par suite, les morphismes f et g seront des isomorphismes réciproques l'un de l'autre. □

Donc, *si l'espace X est séparé, la relation d'ordre $\geqslant$ est définie sur les classes des objets isomorphes de la catégorie* $\mathrm{COV}(X, x_0)$ (quel que soit le point base x_0).

Définition 7. Un revêtement pointé $\pi_0: (\tilde{X}_0, \tilde{x}_0) \to (X, x_0)$ d'un espace séparé pointé (X, x_0) est dit:

a) *universel* si $\pi_0 \geqslant \pi$ pour tout revêtement pointé $\pi: (\tilde{X}, \tilde{x}) \to$ $\to (X, x_0)$;

b) *maximal* si $\pi \geqslant \pi_0$ si et seulement si $\pi_0 \geqslant \pi$, c'est-à-dire si les revêtements π et π_0 sont isomorphes;

c) *simplement connexe* si l'espace $\tilde{X}_0$ l'est aussi.

Il est évident que:

a) *deux revêtements universels quelconques sont équivalents et, par suite, isomorphes*;

b) *tout revêtement universel est maximal*;

c) *si l'espace X est localement connexe, tout revêtement simplement connexe est universel.*

À noter que les réciproques sont généralement fausses: il existe des espaces X localement connexes ayant un revêtement maximal mais pas universel, ou encore universel mais pas simplement connexe. Il existe aussi des espaces dont les revêtements maximaux ne sont pas isomorphes.

Il est évident toutefois que si l'espace X possède un revêtement universel π_0, tout revêtement maximal de X est isomorphe à π_0, de sorte que dans ce cas les revêtements maximaux sont confondus avec les revêtements universels.

En particulier, si un espace localement connexe X possède un revêtement simplement connexe π_0, n'importe lequel de ses revêtements maximaux est isomorphe à π_0, de sorte que les revêtements de tels espaces sont tout à la fois simplement connexes, universels et maximaux.

Définition 8. On dit qu'un espace connexe X est *semi-localement simplement connexe* s'il possède un recouvrement ouvert composé d'ensembles simplement connexes.

Théorème 2. *Tout espace X séparé connexe localement connexe et semi-localement simplement connexe possède un revêtement simplement connexe.*

Corollaire 1. *Les revêtements de tout espace séparé connexe localement connexe et semi-localement simplement connexe sont à la fois simplement connexes, universels et maximaux.* □

Pour prouver le théorème 2, nous aurons besoin d'une construction généralisant celle du coamalgame d'un couple d'applications.

Soit A un ensemble d'indices, et soit donnée pour tout $\alpha \in A$ une application continue (provisoirement arbitraire) $\pi_\alpha : \tilde{X}_\alpha \to X$. Dans le produit $\prod_\alpha \tilde{X}_\alpha$ des espaces $\tilde{X}_\alpha$ considérons le sous-espace $\tilde{X}$ des points $(\tilde{x}_\alpha)$, $\tilde{x}_\alpha \in \tilde{X}_\alpha$, pour lesquels le point $\pi_\alpha(\tilde{x}_\alpha) \in X$ est le même pour tous les α. Sous ces conditions la formule

$$\pi(\{\tilde{x}_\alpha\}) = \pi_\alpha(\tilde{x}_\alpha)$$

définit une application continue

$$\pi : \tilde{X} \to X$$

appelée *coamalgame* des applications π_α. Si les espaces considérés sont pointés et les applications π_α, des applications d'espaces pointés, alors en choisissant pour point base de l'espace $\tilde{X}$ le point $(\tilde{x}_\alpha^{(0)})$, où $\tilde{x}_\alpha^{(0)}$ sont les points bases des espaces $\tilde{X}_\alpha$, on trouve que l'application π est aussi une application d'espaces pointés.

Supposons maintenant que l'espace X est connexe, localement connexe et séparé, et que toutes les applications π_α sont des revêtements (pointés). Dans ces conditions, d'après le corollaire 1 du théorème 1, tout sous-ensemble U ouvert et simplement connexe de l'espace X sera revêtu sans pli par chaque application π_α. Soient $\tilde{U}_{\alpha,\beta_\alpha}$ les composantes de l'ensemble $\pi^{-1}(U_\alpha)$, où β_α parcourt un ensemble d'indices B_α (dépendant de α). Par hypothèse, chaque application

$$\pi_\alpha|_{\tilde{U}_{\alpha,\beta_\alpha}} : \tilde{U}_{\alpha,\beta_\alpha} \to U$$

est un homéomorphisme, et les ensembles $\tilde{U}_{\alpha,\beta_\alpha}$ sont ouverts et disjoints.

Soit $B = \prod_\alpha B_\alpha$ le produit des ensembles B_α. Pour tout $\beta = (\beta_\alpha) \in B$, posons

$$\tilde{U}_\beta = \tilde{X} \cap \prod_\alpha \tilde{U}_{\alpha, \beta_\alpha}.$$

Il est évident que les ensembles $\tilde{U}_\beta \subset \tilde{X}$ sont disjoints et leur somme est l'image réciproque $\pi^{-1}(U)$ de l'ensemble U par l'application π:

$$\pi^{-1}(U) = \bigcup_{\beta \in B} \tilde{U}_\beta.$$

Il est tout aussi évident que la restriction π à tout ensemble $\tilde{U}_\beta$ est un homéomorphisme sur U (l'homéomorphisme réciproque sera l'application $x \mapsto (\sigma_{\alpha, \beta_\alpha}(x))$, où $\sigma_{\alpha, \beta_\alpha}: U \to U_{\alpha, \beta_\alpha}$ est l'homéomorphisme réciproque de $\pi_\alpha|_{U_{\alpha, \beta_\alpha}}$). On remarque, en particulier, que les ensembles $\tilde{U}_\beta$ sont connexes et, par conséquent, sont les composantes de l'ensemble $\pi^{-1}(U)$.

Néanmoins on ne peut affirmer que π revêt sans pli l'ensemble U, puisque les ensembles $\tilde{U}_\beta$ ne sont pas en général des sous-ensembles ouverts de l'espace $\tilde{X}$.

Pour obvier à cet inconvénient, on munit $\tilde{X}$ d'une topologie plus forte (qui possède un plus grand nombre d'ensembles ouverts). Pour définir une topologie sur un ensemble, il suffit manifestement de donner sa base; par ailleurs on sait (ce qui se démontre trivialement) qu'une famille d'ensembles est une base d'une topologie si et seulement si l'intersection $V_1 \cap V_2$ de deux ensembles quelconques de cette famille est la réunion d'ensembles de cette famille. Cette propriété caractéristique des bases est manifestement réalisée pour une famille d'ensembles qui sont les composantes d'ensembles ouverts d'un espace topologique $\tilde{X}$ (si V_1 et V_2 sont les composantes d'ouverts U_1 et U_2, alors $V_1 \cap V_2$ est formée de composantes de l'ouvert $U_1 \cap U_2$). On peut donc prendre cette famille pour base d'une topologie sur $\tilde{X}$. L'espace $\tilde{X}$ muni de cette topologie sera désigné par $\tilde{X}'$, et l'application π, traitée comme une application $\tilde{X}' \to X$, par π'. L'application π' est continue, puisque les ensembles ouverts dans $\tilde{X}$ le sont visiblement dans $\tilde{X}'$.

Les ensembles $\tilde{U}_\beta$ étant les composantes de l'ensemble ouvert $\pi^{-1}(U)$, ils sont ouverts dans $\tilde{X}'$. Mais nous ne pouvons affirmer que l'application π' revêt sans pli l'ensemble X, puisqu'il n'est pas exclu qu'en passant de $\tilde{X}$ à $\tilde{X}'$ l'application π cesse d'être un

homéomorphisme sur $\tilde{U}_\beta$. En fait, l'application π' reste un homéomorphisme sur $\tilde{U}_\beta$. En d'autres termes, la topologie induite sur $\tilde{U}_\beta$ par la topologie de l'espace X' (qui sera appelée topologie II) est confondue avec la topologie initiale induite par la topologie de l'espace $\tilde{X}$ (qui sera appelée topologie I). En effet, pour la topologie I, l'ensemble $\tilde{U}_\beta$, étant homéomorphe à un ensemble ouvert U de l'espace localement connexe X, est lui-même localement connexe, c'est-à-dire possède une base formée d'ensembles connexes. D'autre part, pour la topologie II, tout ensemble ouvert dans $\tilde{U}_\beta$ est de la forme $C \cap \tilde{U}_\beta$, où C est la composante d'un ouvert V de $\tilde{X}$. Soit $\tilde{x}$ un point de $C \cap \tilde{U}_\beta$. L'intersection $V \cap \tilde{U}_\beta$ est un voisinage de ce point pour la topologie I, donc elle contient un voisinage W connexe (donc contenu dans C) du point $\tilde{x}$. Par suite, tout point $\tilde{x} \in C \cap \tilde{U}_\beta$ possède un voisinage W contenu dans $C \cap \tilde{U}_\beta$ pour la topologie I. Cela signifie que l'ensemble $C \cap \tilde{U}_\beta$ est ouvert pour la topologie I. Donc, les topologies I et II sont confondues.

Ceci prouve que l'application $\pi' : \tilde{X}' \to X$ revêt sans pli tout sous-ensemble ouvert simplement connexe $U \subset X$. Donc, si l'espace X est semi-localement simplement connexe, chacun de ses points possède un voisinage revêtu sans pli par l'application π'. Donc, cette application est un revêtement faible, et, par suite, pour toute composante $\tilde{X}_0'$ de l'espace X_0', l'application

$$\pi_0' = \pi|_{\tilde{X}_0'} : \tilde{X}_0' \to X$$

sera un revêtement. Lorsqu'on considère des revêtements pointés, l'arbitraire introduit par le choix de la composante $\tilde{X}_0'$ disparaît, puisque pour une telle composante il faut visiblement prendre une composante contenant le point base $\tilde{x}_0' = (x_\alpha^{(0)}) \in \tilde{X}$.

Donc, si l'espace X est connexe, localement connexe et semi-localement simplement connexe, le procédé exposé permet de construire à l'aide d'une famille quelconque de revêtements pointés $\pi_\alpha : (\tilde{X}_\alpha, \tilde{x}_\alpha^{(0)}) \to (X, x_0)$ un seul revêtement pointé

$$\pi_0' : (\tilde{X}_0', \tilde{x}_0') \to (X, x_0).$$

Par abus de langage, ce revêtement sera appelé *coamalgame des revêtements* π_α.

Il est évident que l'application $f_\alpha : \tilde{X}_0' \to \tilde{X}_\alpha$, définie par la

formule

$$f_\alpha\,((\tilde{x}_\alpha)) = \tilde{x}_\alpha,$$

est continue et vérifie la relation $\pi_0' = \pi_\alpha \circ f_\alpha$, c'est-à-dire est un morphisme du revêtement π_0' dans le revêtement π_α. Donc, *si le revêtement π_0' est coamalgame des revêtements π_α, alors $\pi_0' \geqslant \pi_\alpha$ pour tout $\alpha \in A$.*

Par conséquent, si les revêtements $\pi_\alpha\colon \tilde{X}_\alpha \to X$ forment une *famille complète de revêtements pointés*, c'est-à-dire si tout revêtement pointé $\pi\colon \tilde{X} \to X$ de l'espace X est isomorphe à un revêtement π_{α_0} (l'existence de cette famille est évidente: il suffit de choisir un représentant de chaque classe de revêtements isomorphes de l'espace X), alors le revêtement $\pi_0'\colon X_0' \to X$ est universel. Ceci prouve que *tout espace X connexe localement connexe et semi-localement simplement connexe admet un revêtement universel $\pi_0'\colon \tilde{X}_0' \to X$.*

On prouvera le théorème 2 lorsqu'on aura montré que ce revêtement est simplement connexe pour l'espace séparé X. A cet effet on aura besoin du lemme suivant:

Lemme 4. *Si un espace X connexe localement connexe et semi-localement simplement connexe est séparé, le composé*

$$\pi \circ \rho\colon \tilde{X}_1 \to X$$

de deux revêtements quelconques $\pi\colon \tilde{X} \to X$ et $\rho\colon \tilde{X}_1 \to \tilde{X}$ est aussi un revêtement.

D é m o n s t r a t i o n. Par hypothèse, l'espace X possède un recouvrement formé d'ensembles ouverts simplement connexes U. D'après le corollaire 1 du théorème 1, les ensembles U sont revêtus sans pli par l'application π, et, par suite, les composantes $\tilde{U}$ de leurs images réciproques $\pi^{-1}(U)$ sont des ensembles ouverts homéomorphes aux ensembles U. Donc, l'espace $\tilde{X}$ est semi-localement simplement connexe. D'autre part, il est connexe par hypothèse, et localement connexe et séparé d'après les remarques de cette leçon *in limine*. Donc, les ensembles $\tilde{U}$ sont encore justiciables du corollaire 1 du théorème 1, et, par suite, chacun d'eux est revêtu sans pli par l'application ρ. Mais il est évident alors que l'application $\pi \circ \rho$ revêtira sans pli tous les ensembles U, donc sera un revêtement de l'espace X. □

Nous sommes maintenant en mesure d'achever la démonstration du théorème 2.

Lemme 5. *Si un revêtement universel*

$$\pi_0'\colon \tilde{X}_0' \to X$$

d'un espace X séparé connexe localement connexe et semi-localement simplement connexe est coamalgame d'une famille complète de revêtements $\pi_\alpha: \tilde{X}_\alpha \to X$, il est simplement connexe.

Démonstration. Soit $\rho: \tilde{X}_1 \to \tilde{X}'_0$ un revêtement pointé de l'espace $\tilde{X}'_0$. D'après le lemme 4, le composé $\pi'_0 \circ \rho: \tilde{X}_1 \to X$ est un revêtement, et, par suite, (puisque la famille $\{\pi_\alpha\}$ est complète) est isomorphe à un revêtement $\pi_\alpha: \tilde{X}_{\alpha_0} \to X$. Soit $f: \tilde{X}_{\alpha_0} \to \tilde{X}_1$ l'isomorphisme correspondant. L'application $\rho \circ f: \tilde{X}_{\alpha_0} \to \tilde{X}'_0$ sera alors un morphisme du revêtement π_{α_0} dans le revêtement π'_0 et l'on aura $\pi_{\alpha_0} \geqslant \pi'_0$.

Les revêtements π'_0 et π_{α_0} étant universels, on a $\pi'_0 \geqslant \pi_{\alpha_0}$, ce qui prouve qu'ils sont équivalents et, par suite, isomorphes. Si maintenant $g: \tilde{X}'_0 \to \tilde{X}_{\alpha_0}$ est un isomorphisme du revêtement π'_0 sur le revêtement π_{α_0}, l'application $\sigma = f \circ g: \tilde{X}'_0 \to \tilde{X}_1$ est de toute évidence une section du revêtement ρ.

Donc, le revêtement ρ est trivial d'après la proposition 6.

Par conséquent, tout revêtement de l'espace $\tilde{X}'_0$ est trivial, c'est-à-dire que cet espace est simplement connexe. □

La démonstration du théorème 2 peut prêter à équivoque à cause de la notion de famille complète de revêtements dans la définition de laquelle figurent « tous » les revêtements, ce qui, on le sait, peut conduire à des paradoxes (soit dit en passant, ceci vaut aussi pour la notion d'espace simplement connexe dans la définition duquel participent « tous » ses revêtements, mais pour ne pas interrompre l'exposé nous avons jugé préférable de ne pas nous arrêter sur cette question).

En théorie « naïve » des ensembles, on admet généralement qu'aucun paradoxe n'apparaît si l'on opère seulement avec les sous-ensembles d'un ensemble fixé et avec leurs ensembles quotients. Sans perdre ceci de vue (et en admettant que l'espace pointé donné (X, x_0) est fixe), considérons l'ensemble Σ de toutes les suites finies de la forme

$$(x_1, U_1, \ldots, x_n, U_n), \tag{3}$$

où $x_1, \ldots, x_n$ sont des points de l'espace X, et $U_1, \ldots, U_n$ des ouverts de cet espace tels que pour tout $i = 1, \ldots, n$, les points x_{i-1} et x_i appartiennent à l'ensemble U_i. Dans cet ensemble nous considérons les sous-ensembles Σ' ayant chacun les propriétés suivantes:

1) dans l'espace X il existe un voisinage simplement connexe U du point x_0, tel que $(x_0, U) \in \Sigma'$;

2) il existe un ensemble quotient $\tilde{X} = \Sigma'/\varphi$ de Σ', une topologie sur cet ensemble quotient et une application continue $\pi: \tilde{X} \to X$ tels que

a) π est un revêtement;

b) tous les points de Σ' de la forme (x_0, U), où U est un voisinage simplement connexe du point x_0 dans X, définissent un même point $\tilde{x}_0$ de l'ensemble quotient $\tilde{X}$ et ce point est envoyé dans le point x_0 par l'application π (de sorte que $\pi: (\tilde{X}, \tilde{x}_0) \to (X, x_0)$).

Soient $\pi_\alpha: (\tilde{X}_\alpha, \tilde{x}_\alpha) \to (X, x_0)$ les revêtements pointés de la forme $\Sigma'/\varphi \to X$ obtenus pour tous les choix possibles du sous-ensemble Σ', de la relation d'équivalence φ sur Σ', de la topologie sur Σ'/φ et de l'application π.

Lemme 6. *La famille des revêtements pointés $\pi_\alpha: X_\alpha \to X$ est complète.*

D é m o n s t r a t i o n. Soit $\pi: \tilde{X} \to X$ un revêtement pointé de l'espace X, et soit Σ' le sous-ensemble de Σ des suites (3) pour lesquelles les ensembles $U_1, \ldots, U_n$ sont revêtus sans pli par l'application π. Considérons une suite quelconque (3) de Σ'. Associons au point x_0 le point base $\tilde{x}_0$ de l'espace $\tilde{X}$. Par hypothèse, $\pi(\tilde{x}_0) = x_0$. Procédons par récurrence. Supposons que pour un $i = 1, \ldots, n$ on ait déjà construit un point $\tilde{x}_{i-1} \in X$ tel que $\pi(\tilde{x}_{i-1}) = x_{i-1}$. L'ensemble U_i étant revêtu sans pli par l'application π et $x_{i-1} \in U_i$, il existe dans l'espace $\tilde{X}$ un seul ouvert $\tilde{U}_i$ contenant le point $\tilde{x}_{i-1}$ et homéomorphe à U_{i-1}. Comme $x_i \in U_i$, il existe dans l'ensemble $\tilde{U}_i$ un seul point $\tilde{x}_i$ tel que $\pi(\tilde{x}_i) = x_i$. Nous avons ainsi construit par récurrence tous les points $\tilde{x}_i$, $i = 1, \ldots, n$, et, en particulier, le point $\tilde{x}_n$.

A noter que le point $\tilde{x}_n$ est défini de façon unique par la suite (3). Donc, la formule

$$\varphi(x_1, U_1, \ldots, x_n, U_n) = \tilde{x}_n$$

définit de façon unique une application de l'ensemble Σ' dans l'espace $\tilde{X}$. Si cette application est surjective, l'ensemble quotient Σ'/φ est en correspondance biunivoque avec l'espace $\tilde{X}$. Transportons au moyen de cette correspondance la topologie de l'espace $\tilde{X}$ et l'application π sur l'ensemble quotient Σ'/φ. Il est évident que toutes les conditions posées ci-dessus seront réalisées, de sorte qu'on obtient

un revêtement de la famille $\{\pi_\alpha\}$. Ce revêtement étant par construction isomorphe au revêtement π, on voit que pour achever la démonstration du lemme 6, il reste seulement à prouver que l'application $\varphi: \Sigma' \to \tilde{X}$ est surjective.

Considérons à cet effet l'image $\varphi(\Sigma')$ de l'ensemble Σ' par l'application φ. Si la suite (3) appartient à Σ', en y remplaçant le point x_n par un point quelconque de l'ensemble U_n, on obtient encore une suite de Σ'. Ceci montre que l'ensemble $\varphi(\Sigma')$ contient le point $\tilde{x}_n$ avec son voisinage $\tilde{U}_n$ (cf. plus haut). Donc, l'ensemble $\varphi(\Sigma')$ est ouvert.

Soit maintenant $\tilde{x}$ un point arbitraire de l'adhérence $\overline{\varphi(\Sigma')}$ de l'ensemble $\varphi(\Sigma')$, et soient U un voisinage du point $x = \pi(\tilde{x})$, revêtu sans pli par l'application π, $\tilde{U}$ un voisinage du point $\tilde{x}$, homéomorphe à U. Comme $\tilde{x} \in \overline{\varphi(\Sigma')}$, il vient $\varphi(\Sigma') \cap \tilde{U} \neq \varnothing$. Ceci signifie que dans Σ' il existe une suite (3) telle que $\tilde{x}_n \in \tilde{U}$, et, par suite, $x_n \in U$. Donc, la suite

$$(x_1, U_1, \ldots, x_n, U_n, x, U)$$

appartiendra à Σ' et son image par l'application φ sera le point $\tilde{x}$. Donc, $\tilde{x} \in \varphi(\Sigma')$ et, par suite, l'ensemble $\varphi(\Sigma')$ est fermé.

Etant un sous-ensemble ouvert et fermé (et non vide) de l'espace connexe $\tilde{X}$, l'ensemble $\varphi(\Sigma')$ est confondu avec $\tilde{X}$ tout entier, donc l'application φ est surjective.

Ce qui achève la démonstration du lemme 6. □

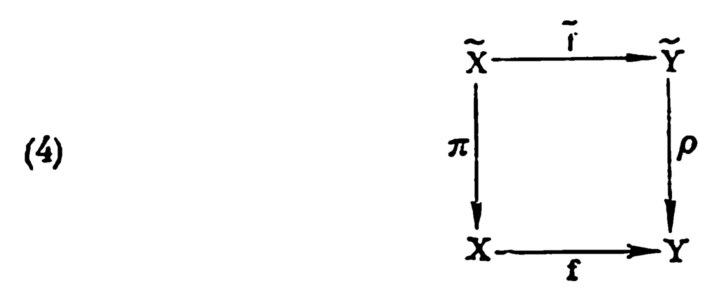
Nous avons par la même occasion entièrement justifié (dans le cadre de la théorie naïve des ensembles) notre construction d'un revêtement universel simplement connexe d'après le principe général indiqué ci-dessus. Que la famille complète construite contienne des revêtements en plusieurs exemplaires (des revêtements isomorphes) ne nuit pas à la cause, car en prouvant le théorème 2 nous n'avons pas exigé que les revêtements figurent en un seul exemplaire. Du reste, on peut aisément éliminer les répétitions des revêtements grâce à l'axiome du choix mais ceci ne fera qu'introduire un arbitraire incontrôlable.

Etudions en conclusion la fonctorialité d'un revêtement universel.

Pour toute catégorie **C** est définie la catégorie AR-**C** dont les objets sont les morphismes $\pi: \tilde{X} \to X$ de la catégorie **C**, et les mor-

phismes, les diagrammes commutatifs de la forme

(4)



où $\tilde{f}$ et f sont les morphismes de la catégorie **C**. Un morphisme (4) est noté ordinairement $(\tilde{f}, f)$: c'est un morphisme d'un objet $\pi: \tilde{X} \to X$ dans un objet $\rho: \tilde{Y} \to Y$. Le composé de ces morphismes se définit de manière évidente.

En particulier, pour $\mathbf{C}$ = TOP et $\mathbf{C}$ = TOP$^{\cdot}$, on obtient les catégories AR-TOP et AR-TOP$^{\cdot}$. On désignera respectivement par COV et COV$^{\cdot}$ les sous-catégories complètes de ces catégories, dont les objets sont des revêtements et on les appellera *catégories de revêtements*. Les morphismes de ces catégories sont des carrés de la forme (4), dans lesquels π et ρ sont des revêtements, et $\tilde{f}$ et f, des applications continues.

Les morphismes de la définition 6 que nous appellerons dorénavant *morphismes au-dessus de* X sont un cas particulier des morphismes $(\tilde{f}, f)$ obtenus pour f = id. Cela signifie que pour *tout espace (pointé)* X, *la catégorie* COV(X) (*resp.* COV$^{\cdot}$(X)) *est une sous-catégorie de la catégorie* COV (*resp.* COV$^{\cdot}$).

En associant à chaque revêtement (pointé) $\pi: \tilde{X} \to X$ l'espace X et à chaque morphisme $(\tilde{f}, f)$ une application f, on obtient de toute évidence un foncteur COV $\to$ TOP (un foncteur COV$^{\cdot}$ $\to$ TOP$^{\cdot}$). Pour tout espace (pointé) X, la catégorie COV(X) (la catégorie COV$^{\cdot}$(X)) est visiblement l'image réciproque de la catégorie triviale (X, id_X) par ce foncteur.

Soit H-TOP$^{\cdot}$ une sous-catégorie complète de TOP$^{\cdot}$, formée des espaces séparés connexes localement connexes et semi-localement simplement connexes, et soit H-COV$^{\cdot}$ son image réciproque par le foncteur COV$^{\cdot}$ $\to$ TOP$^{\cdot}$. En d'autres termes, H-COV$^{\cdot}$ est la sous-catégorie complète de la catégorie COV$^{\cdot}$, composée des revêtements au-dessus d'un espace de H-TOP.

Théorème 3. *Il existe un foncteur*

(5) $$\text{H-TOP}^{\cdot} \to \text{H-COV}^{\cdot},$$

associant à tout espace pointé (X, x_0) *de la catégorie* H-TOP$^{\cdot}$ *un re-*

vêtement pointé universel simplement connexe

$$\pi_X : (\tilde{X}, \tilde{x}_0) \to (X, x_0). \qquad (6)$$

Ce foncteur est unique à un isomorphisme près.

Démonstration. Définissons le foncteur (5) sur des objets en choisissant arbitrairement un revêtement (6) pour chaque espace (X, x_0). Soit $f: (X, x_0) \to (Y, y_0)$ une application pointée. L'espace $\tilde{X}$ étant simplement connexe, le théorème 1 nous dit qu'il existe une seule application $\tilde{f}: (\tilde{X}, \tilde{x}_0) \to (\tilde{Y}, \tilde{y}_0)$ qui est un relèvement de l'application $\pi_X \circ f$ relativement au revêtement π_Y et qui forme avec f le morphisme $(\tilde{f}, f)$ du revêtement π_X dans le revêtement π_Y. En associant à l'application f le morphisme $(\tilde{f}, f)$, on obtient (en vertu de l'unicité de l'application $\tilde{f}$) le foncteur cherché (5).

En choisissant un autre revêtement (6), on obtient un foncteur isomorphe (le fait que l'isomorphisme sera fonctoriel résulte encore de l'unicité). □

Remarque 1. La construction du foncteur (5) donne lieu à un arbitraire assez désagréable. Bien que cet arbitraire ne change rien au fond du problème en conduisant à des foncteurs canoniquement isomorphes, on peut s'en débarrasser au besoin en prenant pour (6) le coamalgame de la famille complète des revêtements du lemme 6. (A noter que la construction correspondante définit non seulement un seul espace $\tilde{X}$, mais aussi un seul point base $\tilde{x}_0$.)

LEÇON 9

Revêtements différentiables.— Isomorphisme des catégories des revêtements différentiables et topologiques.— Existence de revêtements différentiables universels.— Revêtements de groupes différentiables et topologiques.— Revêtements universels de groupes de Lie.— Lemmes relatifs aux groupes topologiques.— Isomorphismes locaux et revêtements.— Description de groupes de Lie localement isomorphes.

Appliquons les résultats de la leçon précédente aux variétés différentiables et aux groupes de Lie. Etant donnée une variété différentiable (connexe) M, on appellera *revêtements topologiques* ses revêtements $\widetilde{M} \to M$ au sens de la définition 1 de la leçon 8, c'est-à-dire ses revêtements pour sa structure d'espace topologique, et on désignera par $\mathrm{COV}_{\mathrm{top}}(M)$ la catégorie de ces revêtements.

Définition 1. On dit qu'un revêtement $\pi: \widetilde{M} \to M$, où $\widetilde{M}$ et M sont des variétés différentiables connexes, est *différentiable* s'il est une application différentiable de $\widetilde{M}$ sur M et si pour tout ouvert connexe $U \subset M$ revêtu sans pli par l'application π, la restriction de l'application π à chaque composante $\widetilde{U}$ de l'ensemble $\pi^{-1}(U)$ est un difféomorphisme $\widetilde{U} \to U$.

La dernière condition signifie, en particulier, que tout revêtement différentiable est étale (est un difféomorphisme local).

Le revêtement trivial $\mathbb{R} \to \mathbb{R}$, $x \to x^3$ est un exemple de revêtement non différentiable qui est une application différentiable.

On désignera par COV-DIFF la catégorie dont les objets sont les revêtements différentiables des variétés différentiables, et les morphismes. leurs morphismes $(\widetilde{f}, f)$ en tant que revêtements topologiques,

où $\tilde{f}$ et f sont les applications différentiables, et par COV(M), la sous-catégorie de COV-DIFF formée par les revêtements de M et leurs morphismes au-dessus de M. Par COV$^{\bullet}$-DIFF et COV$^{\bullet}$(M) on désignera les variantes pointées de ces catégories.

L'annihilation de la structure différentiable définit pour toute variété différentiable M un foncteur

$$(1) \qquad \mathrm{COV}(M) \to \mathrm{COV}_{\mathrm{top}}(M).$$

Proposition 1. *Le foncteur* (1) *est un isomorphisme de catégories, c'est-à-dire est une application bijective sur les objets et les morphismes.*

Démonstration. La bijectivité du foncteur (1) sur les objets signifie que pour tout revêtement topologique $\pi: \tilde{M} \to M$, il existe sur $\tilde{M}$ une seule structure différentiable pour laquelle π est un revêtement différentiable. Pour le prouver considérons une carte (U, h) de M pour laquelle U est connexe et revêtu sans pli par l'application π. Soit $\tilde{U}$ une composante arbitraire de l'ensemble $\pi^{-1}(U)$, et soit $\tilde{h} = h \circ (\pi |_{\tilde{U}})$. Il est évident que le couple $(\tilde{U}, \tilde{h})$ est une carte de $\tilde{M}$. Si $(\tilde{V}, \tilde{k})$ est une telle carte, alors $(\tilde{U}, \tilde{h})$ et $(\tilde{V}, \tilde{k})$ sont compatibles, puisque $\tilde{h} \circ \tilde{k}^{-1} = h \circ k^{-1}$ sur l'ensemble $\tilde{k}(\tilde{U} \cap \tilde{V}) = k(U \cap V)$. Ceci montre que les cartes de la forme $(\tilde{U}, \tilde{h})$ constituent un atlas de $\tilde{M}$. L'application π sera visiblement un revêtement différentiable pour la structure différentiable correspondante sur $\tilde{M}$. L'unicité de cette structure résulte directement du fait que les couples $(\tilde{U}, \tilde{h})$ sont des cartes différentiables pour toute structure différentiable sur $\tilde{M}$ pour laquelle π est un revêtement différentiable.

La bijectivité du foncteur (1) sur les morphismes signifie désormais que chaque morphisme $\tilde{f}: \tilde{M}_1 \to \tilde{M}$ des revêtements différentiables $\pi_1: \tilde{M}_1 \to M$ et $\pi: \tilde{M} \to M$ traités comme des revêtements topologiques est une application différentiable. Or, ceci est évident d'après ce qui a été dit plus haut. En effet, l'application $\tilde{f}$ étant un revêtement (lemme 3 de la leçon précédente), il existe un atlas de M formé de cartes (U, h) revêtues sans pli par l'application π, telles que toutes les cartes correspondantes $(\tilde{U}, \tilde{h})$ de $\tilde{M}$ sont revêtues sans pli par l'application $\tilde{f}$. Ceci signifie que pour toute carte $(\tilde{U}, \tilde{h})$, il existe une carte $(\tilde{U}_1, \tilde{h}_1)$ de $\tilde{M}_1$ revêtant la carte (U, h) et telle que f est un homéomorphisme de $\tilde{U}_1$ sur $\tilde{U}$.

Mais alors $\tilde{h}_1 = \tilde{h} \circ \tilde{f}$, et, par suite, l'application $\tilde{f}$ est définie sur ces cartes par l'application identique (donc différentiable). Par conséquent, l'application $\tilde{f}$ est différentiable. □

Remarque 1. Le foncteur d'annihilation

$$\text{COV-DIFF} \to \text{COV}$$

n'est pas un isomorphisme de catégories. Sur les objets ce foncteur ne sera ni surjectif (en raison de l'existence d'espaces topologiques non localement euclidiens), ni injectif (en raison de la possibilité de munir un même espace topologique de structures différentiables différentes), et son image ne sera pas une sous-catégorie complète de la catégorie COV. On peut tout au plus affirmer (à titre de généralisation directe de la partie de la proposition 1 relative aux morphismes) que *si pour un morphisme* $(\tilde{f}, f)$ *d'un revêtement différentiable* $\pi_1 : \tilde{M}_1 \to M_1$ *dans un revêtement différentiable* $\pi: \tilde{M} \to M$ *l'application* f *est différentiable, il en est de même de l'application* $\tilde{f}$. En effet, (cf. démonstration de la proposition 1), pour deux cartes quelconques (U_1, h_1) et (U, h) des variétés M_1 et M, l'application $\tilde{f}$ est définie sur les cartes $(\tilde{U}_1, \tilde{h}_1)$ et $(\tilde{U}, \tilde{h})$ revêtant respectivement (U_1, h_1) et (U, h) par les mêmes fonctions (donc différentiables) que l'application f sur les cartes (U_1, h_1) et (U, h) (on admet bien sûr que $\tilde{f}(\tilde{U}_1) \subset \tilde{U}$, et, par suite, $f(U_1) \subset U$). □

Etant un espace topologique localement euclidien, toute variété différentiable est localement connexe. Il s'avère qu'elle est en plus semi-localement simplement connexe et même *localement simplement connexe*, c'est-à-dire possède une base formée d'ensembles ouverts simplement connexes. Ceci résulte immédiatement du lemme suivant:

Lemme 1. *Le cube unité* Q *de l'espace* $\mathbb{R}^n$ *(qu'il soit ouvert ou fermé) est un espace simplement connexe.*

Prouvons tout d'abord le lemme suivant:

Lemme 2. *Si des espaces* X *et* Y *simplement connexes (et connexes) sont séparés et localement connexes, leur produit* $X \times Y$ *est aussi simplement connexe.*

Démonstration. Soit $\pi: \tilde{Z} \to X \times Y$ un revêtement de l'espace $X \times Y$. Il nous faut prouver que ce revêtement est trivial, c'est-à-dire (cf. proposition 4 de la leçon 8) qu'il possède une section $\sigma: X \times Y \to \tilde{Z}$.

Choisissons (et figeons) des points $x_0 \in X$, $y_0 \in Y$ et $\tilde{z}_0 \in \in \pi^{-1}(x_0, y_0)$. D'après le théorème 1, il existe une seule application

$$\tau: (Y, y_0) \to (\tilde{Z}, \tilde{z}_0)$$

telle que $(\pi \circ \tau)(y) = (x_0, y)$ pour tout point $y \in Y$ (cette application n'est autre que le relèvement de l'application $y \mapsto (x_0, y)$ de l'espace pointé connexe localement connexe et simplement connexe (Y, y_0) dans l'espace pointé séparé $(X \times Y, (x_0, y_0))$). De façon analogue, pour tout point $y \in Y$, il existe une seule application

$$\sigma_y: (X, x_0) \to (\tilde{Z}, \tau(y)),$$

telle que

$$(\pi \circ \sigma_y)(x) = (x, y)$$

pour tout point $x \in X$. Mais alors l'application

$$\sigma: X \times Y \to \tilde{Z},$$

définie par la formule

$$\sigma(x, y) = \sigma_y(x), \quad (x, y) \in X \times Y$$

satisfera de toute évidence à la relation $\pi \circ \sigma = \mathrm{id}$. Donc, pour achever la démonstration, il nous reste à prouver seulement que l'application σ est continue.

Pour cela, il suffit seulement de montrer que l'image $A = \sigma(X \times Y)$ de l'espace $X \times Y$ par l'application σ est un sous-ensemble ouvert de l'espace $\tilde{Z}$. En effet, tout point $\tilde{z} \in A$ aura alors un voisinage W entièrement contenu dans A, et l'image $\pi(W)$ de ce voisinage W par l'application π sera un voisinage du point $\pi(\tilde{z})$ appliqué sur W au moyen de σ.

Chosissons un point arbitraire $y_1 \in Y$ et considérons le sous-ensemble $\sigma_{y_1}(X)$ de l'ensemble $\tilde{Z}$ (sous-ensemble qui, on le sait, est homéomorphe à l'espace X). Soit B l'ensemble de tous les points intérieurs de A contenus dans $\sigma_{y_1}(X)$. Il nous faut prouver que $B = \sigma_{y_1}(X)$. L'ensemble B étant ouvert dans $\sigma_{y_1}(X)$, il nous suffit de montrer qu'il n'est pas vide et qu'il est fermé dans $\sigma_{y_1}(X)$.

Soit $\tilde{z}_1 \in \sigma_y(X)$, et soit $\pi(\tilde{z}_1) = (x_1, y_1)$. Dans X et Y, il existe des voisinages connexes U et V de x_1 et y_1 tels que $U \times V$ est revêtu sans pli par l'application π. Soit W la composante du point $\tilde{z}_1$ dans $\pi^{-1}(U \times V)$. Il est évident que

$$W = \bigcup_{y \in V} \tilde{U}(y),$$

où $\tilde{U}(y) = W \cap \pi^{-1}(U \times \{y\})$, et de plus chaque ensemble $\tilde{U}(y)$ soit ne coupe pas A, soit y est entièrement contenu.

Si $x_1 = x_0$, les ensembles $\widetilde{U}(y)$ ne sont autres que les intersections de W et des ensembles $\sigma_y(X)$:

$$\widetilde{U}(y) = W \cap \sigma_y(X).$$

Donc, dans ce cas $\widetilde{U}(y) \subset A$ et, par suite, $W \subset A$. Comme W est ouvert dans $\widetilde{Z}$, cela signifie que le point $\widetilde{z}_1$, tel que $x_1 = x_0$, appartient à B. Par conséquent, l'ensemble B n'est pas vide.

Supposons maintenant que le point $\widetilde{z}_1$ appartient à l'adhérence $\overline{B}$ de B. Alors, l'ensemble $\widetilde{U}(y_1)$ (qui est un voisinage du point $\widetilde{z}_1$ dans $\sigma_{y_1}(X)$) coupe B. Soit $\widetilde{z}_2 \in B \cap \widetilde{U}(y_1)$, et soit $\pi(\widetilde{z}_2) = (x_2, y_2)$. Le point $\widetilde{z}_2$ étant un point intérieur de l'ensemble A, il existe dans l'espace Y un voisinage V' du point y_1 tel que $V' \subset V$ et $W \cap \pi^{-1}(\{x_2\} \times V') \subset A$. Cela signifie que pour tout point $y \in V'$, l'ensemble $\widetilde{U}(y)$ coupe A. On a alors nécessairement $\widetilde{U}(y) \subset A$. Donc l'ensemble

$$W' = \bigcup_{y \in V'} \widetilde{U}(y)$$

est contenu dans A. L'ensemble W' est visiblement ouvert dans Z' et contient le point $\widetilde{z}_1$. Donc, $\widetilde{z}_1 \in B$ et l'ensemble B est fermé. □

Nous pouvons maintenant prouver le lemme 1.

D é m o n s t r a t i o n d u l e m m e 1. D'après le lemme 2, il suffit de traiter le cas où le cube Q est de dimension un, c'est-à-dire est un intervalle de $\mathbb{R}$. Supposons d'abord que cet intervalle est fermé. Alors, pour tout revêtement $\pi: \widetilde{Q} \to Q$ il existe un recouvrement fini de Q dont les éléments sont des intervalles $I_1, \ldots, I_n$ ouverts (dans Q) revêtus sans pli par π. Il est clair que ces intervalles peuvent être numérotés successivement, c'est-à-dire de telle sorte que pour tout $k = 1, \ldots, n$, la réunion J_k des intervalles $I_1, \ldots$ $\ldots, I_k$ soit connexe, et, par suite, soit aussi un intervalle. Les intersections $J_k \cap I_{k+1}$ seront alors connexes, et, par suite, en appliquant $n - 1$ fois le corollaire 1 de la proposition 1 de la leçon précédente, on trouve que l'intervalle $J_n = Q$ est revêtu sans pli par l'application π. Donc, le revêtement π est trivial, et, par suite, Q est simplement connexe.

L'intervalle ouvert Q est la réunion d'une suite croissante d'intervalles fermés au-dessus de chacun desquels le revêtement π est trivial, c'est-à-dire est un homéomorphisme. Donc, π sera un homéomorphisme au-dessus de Q tout entier. Par conséquent, l'intervalle ouvert Q est aussi simplement connexe. □

Corollaire. *La sphère S^n est simplement connexe pour $n \geqslant 2$.*

D é m o n s t r a t i o n. Soient p et q deux points diamétralement opposés de S^n, et soient $U = S^n \setminus \{p\}$ et $V = S^n \setminus \{q\}$. Les ensembles ouverts U et V sont homéomorphes à un cube ouvert à n dimensions, et donc, sont simplement connexes. Leur intersection $U \cap V = S^n \setminus (\{p\} \cup \{q\})$ est homéomorphe au produit $S^{n-1} \times]0, 1[$ de l'équateur S^{n-1} de la sphère S^n par l'intervalle ouvert $]0, 1[$ et, par suite, est connexe (pour $n - 1 \geqslant 1$). Donc, la sphère $S^n = U \cup V$ est simplement connexe d'après le corollaire 2 du théorème 1 de la leçon 8. □

Soit H-DIFF la catégorie des variétés séparées différentiables pointées. D'après ce qui a été prouvé, le foncteur d'annihilation de la structure différentiable envoie cette catégorie dans la catégorie H-TOP$^\bullet$. En combinant ce foncteur à celui du théorème 3 de la leçon précédente, on associe à toute variété M un revêtement simplement connexe $\tilde{M} \to M$. D'après la proposition 1, ce revêtement est muni d'une seule structure de revêtement différentiable. Il est évident que l'on obtient un foncteur de la catégorie H-DIFF$^\bullet$ dans la catégorie H-COV$^\bullet$ des revêtements pointés au-dessus des variétés séparées.

Enonçons ce fait sous la forme du théorème suivant:

Théorème 1. *Il existe un foncteur*

$$\text{H-DIFF}^\bullet \to \text{H-COV}^\bullet,$$

associant à toute variété séparée pointée M un revêtement pointé universel simplement connexe de M

$$\pi_M: \tilde{M} \to M.$$

Ce foncteur est unique à un isomorphisme près. □

Passons maintenant aux groupes de Lie. La notion de revêtement pour les groupes de Lie s'introduit par la définition suivante :

Définition 2. Soient $\tilde{G}$ et G des groupes de Lie connexes. On dit qu'un revêtement différentiable $\pi: \tilde{G} \to G$ est un *revêtement de groupe* s'il est un homomorphisme de groupes.

On introduit de façon analogue la notion de revêtement de groupe pour le cas où $\tilde{G}$ et G sont des groupes topologiques connexes.

On appelle morphismes de revêtements de groupe les morphismes $(\tilde{f}, f)$ de ces revêtements traités comme des revêtements différentiables (ou topologiques), pour lesquels les applications $\tilde{f}$ et f sont des homomorphismes.

Etant donné que dans chaque groupe le point base l'est de façon naturelle — c'est l'unité de ce groupe — et que tout homomorphis-

me envoie l'unité dans l'unité, tous les revêtements de groupe et tous leurs morphismes sont automatiquement pointés.

Les revêtements de groupe des groupes de Lie (ou des groupes topologiques) et leurs morphismes forment la catégorie COV-GR.

Tout groupe de Lie G définit une sous-catégorie de la catégorie COV-GR dont les objets sont les revêtements de groupe du groupe G et les morphismes, les morphismes de COV-GR de la forme $(\tilde{f}, \mathrm{id})$. Sans crainte d'ambiguïté nous désignerons cette sous-catégorie par COV(G) (et la catégorie des revêtements du groupe G traité comme une variété séparée pointée, par $\mathrm{COV}_{\mathrm{diff}}(G)$).

Il s'avère qu'à l'instar du foncteur (1) *le foncteur d'annihilation*

$$\mathrm{GOV}\,(G) \to \mathrm{GOV}_{\mathrm{diff}}\,(G) \tag{2}$$

est un isomorphisme de catégories.

Nous ne prouverons cette proposition que partiellement.

Pour tout revêtement de groupe $\pi\colon \tilde{G} \to G$ les applications

$$\begin{aligned} \tilde{\mu}\colon \tilde{G} \times \tilde{G} \to \tilde{G}, &\qquad (\tilde{x}, \tilde{y}) \mapsto \widetilde{xy}, \\ \tilde{\nu}\colon \tilde{G} \to \tilde{G}, &\qquad \tilde{x} \mapsto \tilde{x}^{-1}, \end{aligned} \tag{3}$$

sont des relèvements (par rapport à π) des applications

$$\begin{aligned} \mu\colon \tilde{G} \times \tilde{G} \to G, &\qquad (\tilde{x}, \tilde{y}) \mapsto \pi\tilde{x}\cdot\pi\tilde{y}, \\ \nu\colon \tilde{G} \to G, &\qquad \tilde{x} \to (\pi\tilde{x})^{-1}. \end{aligned}$$

Ces relèvements étant caractérisés de façon unique (proposition 3 de la leçon précédente) par les conditions

$$\tilde{\mu}\,(\tilde{e}, \tilde{e}) = \tilde{e}, \quad \tilde{\nu}\,(\tilde{e}) = \tilde{e}, \tag{4}$$

où $\tilde{e}$ est l'unité du groupe $\tilde{G}$, le foncteur (2) est une application injective sur les objets.

Pour montrer qu'il est bijectif, il faut pour tout revêtement pointé $\pi\colon (\tilde{G}, \tilde{e}) \to (G_{\mathrm{diff}}, e)$, où G_{diff} est le groupe G traité comme une variété différentiable, et e son unité, montrer que la variété $\tilde{G}$ peut être munie d'une structure de groupe de Lie pour laquelle π sera un revêtement de groupe. Il est clair que pour cela il faut considérer des applications μ et ν (dont la construction n'implique pas que $\tilde{G}$ soit un groupe) et leurs relèvements $\tilde{\mu}$ et $\tilde{\nu}$ qui vérifient les relations (4). *Supposons que les relèvements $\tilde{\mu}$ et $\tilde{\nu}$ existent.*

Les applications $\tilde{\mu} = \pi \circ \tilde{\mu}$ et $\nu = \pi \circ \tilde{\nu}$ étant différentiables et π étant un difféomorphisme local, les applications $\tilde{\mu}$ et $\tilde{\nu}$ sont aussi

différentiables. Par ailleurs, il est évident que l'application π est un homomorphisme relativement aux opérations

$$\widetilde{xy} = \tilde{\mu}(\tilde{x}, \tilde{y}) \quad \text{et} \quad \tilde{x}^{-1} = \tilde{\nu}(\tilde{x}).$$

Donc, l'application π sera un revêtement de groupe lorsqu'on aura prouvé que ces opérations satisfont les axiomes de groupe.

Comme $(\pi \circ \tilde{\mu})(\tilde{x}, \tilde{e}) = \pi(\tilde{x})$ et $(\pi \circ \tilde{\mu})(\tilde{e}, \tilde{x}) = \pi(\tilde{x})$, les applications $\tilde{x} \mapsto \tilde{\mu}(\tilde{e}, \tilde{x})$ et $\tilde{x} \mapsto \tilde{\mu}(\tilde{x}, \tilde{e})$ sont des relèvements de l'application $\pi: \tilde{G} \to G$, ayant la propriété suivante: $\tilde{e} \mapsto \tilde{e}$. Cette propriété est également caractéristique de l'application identique id: $\tilde{G} \to \tilde{G}$ (qui est aussi un relèvement de l'application π). Donc, en vertu de l'unicité des relèvements (proposition 3 de la leçon 8), on a les égalités $\tilde{\mu}(\tilde{x}, \tilde{e}) = \tilde{\mu}(\tilde{e}, \tilde{x}) = \tilde{x}$, c'est-à-dire les égalités $\widetilde{xe} = \widetilde{ex} = \tilde{x}$ qui expriment que le point $\tilde{e}$ est l'unité de la multiplication μ.

De façon analogue, les applications $(\tilde{x}, \tilde{y}, \tilde{z}) \mapsto (\widetilde{xy})\tilde{z}$ et $(\tilde{x}, \tilde{y}, \tilde{z}) \mapsto \tilde{x}(\widetilde{yz})$ sont des relèvements de la même application $(\tilde{x}, \tilde{y}, \tilde{z}) \mapsto \pi\tilde{x}\cdot\pi\tilde{y}\cdot\pi\tilde{z}$, le point $(\tilde{e}, \tilde{e}, \tilde{e})$ étant envoyé dans le même point $\tilde{e}$ par ces deux applications. Donc, $(\widetilde{xy})\tilde{z} = \tilde{x}(\widetilde{yz})$ pour tous éléments $\tilde{x}, \tilde{y}, \tilde{z} \in \tilde{G}$, de sorte que la multiplication μ est associative.

Enfin, l'application $\tilde{x} \mapsto \widetilde{xx^{-1}}$ est une application continue de l'espace connexe $\tilde{G}$ dans l'espace discret $\pi^{-1}(e)$ envoyant $\tilde{e}$ dans $\tilde{e}$. Donc, $\widetilde{xx^{-1}} = \tilde{e}$ pour tout $\tilde{x} \in \tilde{G}$.

Par conséquent, $\tilde{G}$ est un groupe. □

Ainsi, la surjectivité du foncteur (2) sur les objets est subordonnée à l'existence des relèvements (3). Nous ne prouverons pas leur existence en général, mais seulement dans le cas où la variété $\tilde{G}$ est simplement connexe.

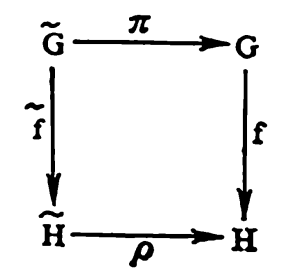
Dans ce cas, l'existence des relèvements $\tilde{\mu}$ et $\tilde{\nu}$ est affirmée par le théorème 1 de la leçon précédente, puisque d'après le lemme 2, le produit $\tilde{G} \times \tilde{G}$ de la variété simplement connexe $\tilde{G}$ par elle-même est également simplement connexe.

Nous avons ainsi prouvé la proposition suivante:

Proposition 2. *Soit G un groupe de Lie, et soit $\pi: \tilde{G} \to G$ un revêtement différentiable simplement connexe de G traité comme une variété pointée différentiable. Sous ces conditions, on peut munir la variété différentiable $\tilde{G}$ d'une seule multiplication qui fait d'elle un*

groupe de Lie, et de l'application π, un homomorphisme (donc, un revêtement de groupe). □

Affirmer que le foncteur (2) est bijectif sur les morphismes revient à affirmer que pour tout couple $\pi: \tilde{G} \to G$ et $\pi_1: \tilde{G}_1 \to G$ de revêtements de groupe, tout morphisme $\tilde{f}: \tilde{G} \to \tilde{G}_1$ de ces revêtements traités comme des revêtements différentiables est un homomorphisme $\tilde{G} \to \tilde{G}_1$. Nous établirons un résultat plus général relatif aux revêtements $\pi: \tilde{G} \to G$ et $\rho: \tilde{H} \to H$ de deux groupes de Lie G et H généralement distincts et à un morphisme $(\tilde{f}, f)$ de la catégorie COV$^{\cdot}$-DIFF du revêtement π dans le revêtement ρ:



Si f est un homomorphisme de groupes, il en est de même de $\tilde{f}$. En effet, les applications $\tilde{G} \times \tilde{G} \to \tilde{H}$ définies par les formules $(\tilde{x}, \tilde{y}) \mapsto \tilde{f}(\tilde{x})\,\tilde{f}(\tilde{y})$ et $(\tilde{x}, \tilde{y}) \mapsto \tilde{f}(xy)$, sont toutes deux des relèvements d'une même application

$$(\tilde{x}, \tilde{y}) \mapsto f(x)\,f(y) = f(xy), \quad \text{où } x = \pi(\tilde{x}), \quad y = \pi(\tilde{y})$$

et, par suite, sont confondues. □

Il s'ensuit de là que le théorème 1 est valable pour les groupes de Lie :

Théorème 2. *Il existe un foncteur* GR-DIFF → GR-DIFF *associant à tout groupe de Lie connexe G un revêtement de groupe simplement connexe de G*

$$\pi_G : \tilde{G} \to G.$$

Ce foncteur est unique à un isomorphisme près.

Le noyau Ker π_G du revêtement $\pi_G: \tilde{G} \to G$ est défini de façon unique (à un isomorphisme près) par le groupe de Lie G. Ce noyau s'appelle *groupe fondamental* (ou *groupe de Poincaré*) du groupe G et se note $\pi_1 G$ (on démontre que le groupe $\pi_1 G$ est confondu avec le groupe fondamental — connu en topologie — $\pi_1 G_{top}$, où G_{top} est le groupe G, traité comme un espace topologique; cf. [8]).

Pour les applications que nous envisageons, il est commode d'énoncer le théorème 2 sous une forme plus algébrique. A cet effet, nous aurons besoin de quelques lemmes simples sur les groupes topologiques.

Lemme 3. *Tout sous-groupe ouvert H d'un groupe topologique G est fermé.*

Démonstration. Le sous-groupe H étant ouvert, toute classe Hx à droite suivant H est aussi ouverte. La réunion d'une famille quelconque de ces classes est donc un ensemble ouvert. En particulier, la réunion de toutes les classes Hx distinctes du sous-groupe H est un ouvert. Or, cette réunion est le complémentaire du sous-groupe H dans G. Donc, le sous-groupe H est fermé. □

Lemme 4. *Tout voisinage V de l'unité d'un groupe topologique connexe G engendre le groupe G.*

Démonstration. Soit H un sous-groupe engendré par le voisinage V. Comme $Vx \subset H$ pour tout $x \in H$, le sous-groupe H est ouvert. Le sous-groupe H est alors aussi fermé d'après le lemme 3. Donc, $H = G$, puisque G est connexe par hypothèse. □

On rappelle que pour tout sous-groupe invariant K d'un groupe topologique G, le groupe quotient G/K est muni d'une *topologie quotient* pour laquelle l'ensemble $V \subset G/K$ est ouvert si et seulement si l'est son image réciproque $\pi^{-1}(V)$ par l'épimorphisme canonique $\pi: G \twoheadrightarrow G/K$. Le groupe quotient G/K est un groupe topologique pour cette topologie.

Lemme 5. *Pour tout sous-groupe invariant $K \subset G$, l'épimorphisme canonique $\pi: G \twoheadrightarrow G/K$ est une application ouverte. Pour tout épimorphisme ouvert $\Phi: G \twoheadrightarrow H$, le groupe quotient G/K du groupe G par le noyau $K = \operatorname{Ker} \Phi$ de l'épimorphisme Φ est isomorphe au groupe H. L'isomorphisme $\varphi: G/K \twoheadrightarrow H$ peut être choisi de telle sorte que l'on ait le diagramme commutatif*

(5)


Démonstration. Si U est ouvert dans G, l'ensemble UK l'est aussi en tant que réunion d'ensembles ouverts Ux, $x \in K$. Or, il est évident que $UK = \pi^{-1}(\pi U)$. Donc, l'ensemble πU est ouvert dans G/K. Ceci prouve la première affirmation.

Il est évident que la formule $\varphi(xK) = \Phi(x)$ définit un seul isomorphisme algébrique $\varphi: G/K \to H$ rendant le diagramme (5) commutatif. Il nous faut donc établir seulement que φ est un homéomorphisme.

Mais si V est ouvert dans H, il en est de même de $\Phi^{-1}(V)$ dans G (puisque Φ est continu), et comme π est ouvert, il en est de même de $\pi(\Phi^{-1}(V)) = \varphi^{-1}(V)$ dans G/K. De façon analogue, si W est ouvert dans G/K, $\pi^{-1}(W)$ l'est aussi dans G, et comme Φ est ouvert par hypothèse, $\Phi(\pi^{-1}(W)) = \varphi(W)$ l'est aussi dans H. Donc, φ est un homéomorphisme. □

Définition 3. Soient G et H des groupes topologiques (ou différentiables) connexes. On dira qu'un homomorphisme $\Phi: G \to H$ du groupe G dans le groupe H est un *isomorphisme local* s'il applique bijectivement un voisinage U de l'unité de G sur un voisinage V de l'unité de H.

Le groupe H étant engendré par le voisinage V en vertu du lemme 4, tout isomorphisme local Φ est un épimorphisme, et comme cet épimorphisme est ouvert, car difféomorphisme local, on peut lui appliquer le lemme 5. Donc, le groupe H est isomorphe au groupe quotient G/K, où $K = \operatorname{Ker} \Phi$. Ceci étant, par hypothèse $K \cap U =$ $= \{e\}$, ce qui signifie par définition que le sous-groupe K est un *sous-groupe discret* du groupe G. Réciproquement, si K est un sous-groupe discret invariant du groupe G, l'épimorphisme canonique $\pi: G \to G/K$ est de toute évidence isomorphisme local. D'après le lemme 5, ceci prouve qu'*il existe un isomorphisme local $G \to H$ si et seulement si le groupe H est isomorphe au groupe quotient G/K du groupe G par un sous-groupe discret invariant K.*

Tout revêtement de groupe $G \to H$ est visiblement un isomorphisme local. Réciproquement, tout isomorphisme local $\Phi: G \to H$ est un revêtement de groupe, puisqu'il revêt sans pli le voisinage V de la définition 3 (en vertu du lemme 5, on peut, sans nuire à la généralité, admettre que $H = G/K$, où $K = \operatorname{Ker} \Phi$ et que Φ est l'épimorphisme canonique $\pi: G \to G/K$; sous ces conditions

$$\Phi^{-1}(V) = UK = \bigcup_{x \in K} Ux,$$

où les ensembles ouverts $Ux \subset G$ sont disjoints et homéomorphes à U) et, par conséquent, le voisinage Vh de tout élément $h \in H$. Donc, *les revêtements de groupe et les isomorphismes locaux représentent la même chose.*

Il est évident que la dernière affirmation est valable pour les groupes de Lie aussi. Bien plus, il est immédiat de voir que si est donné un revêtement de groupe (= isomorphisme local) $G \to H$, où H est un groupe topologique et G un groupe de Lie, alors H est muni d'une seule structure différentiable qui fait de H un groupe

de Lie et de l'application $G \to H$ un revêtement différentiable. En particulier, pour tout sous-groupe discret invariant K du groupe de Lie G, le groupe quotient G/K est aussi un groupe de Lie et, par suite, pour tout isomorphisme local $\Phi: G \to H$ de groupes de Lie, le diagramme commutatif (5) est un diagramme au-dessus de la catégorie des groupes de Lie.

Tout ceci signifie que les groupes de Lie (de même que les groupes topologiques) sont justiciables de la proposition suivante:

Proposition 3. *Une application $\Phi: G \to H$ de groupes connexes est un revêtement de groupe si et seulement si*

a) *elle est un isomorphisme local ou, ce qui est équivalent, si*

b) *l'on a le diagramme commutatif* (5), *où K est un sous-groupe discret invariant, π un épimorphisme canonique, φ un isomorphisme.* □

Cette proposition donne un relief particulier au lemme assez insolite suivant:

Lemme 6. *Tout sous-groupe discret invariant K d'un groupe topologique (et, notamment, différentiable) connexe G appartient au centre de ce groupe (et, notamment, est abélien).*

Démonstration. Soit $x \in K$, et soit U un voisinage de x ne contenant pas d'autres éléments de K. Considérons un voisinage V de l'unité du groupe G tel que $VxV^{-1} \subset U$. (L'existence de ce voisinage résulte immédiatement de la continuité de l'application $y \mapsto yxy^{-1}$.) Le sous-groupe K étant invariant et $U \cap K = \{e\}$, on a $yxy^{-1} = x$ pour tout élément $y \in V$. Ceci signifie que le *centralisateur de x* (le sous-groupe de tous les éléments permutables à x) contient le voisinage V. Donc, ce centralisateur est confondu avec G, puisque V engendre G, et, par suite, x appartient au centre du groupe G. □

Corollaire. *Le groupe fondamental $\pi_1 G$ de tout groupe de Lie connexe G est un groupe abélien.* □

Nous sommes en mesure de formuler notre théorème final. Dans ce théorème, on dira que des groupes de Lie connexes G et H sont *localement isomorphes* s'il existe des isomorphismes locaux de la forme $P \to G$ et $P \to H$, où P est un groupe de Lie connexe.

Théorème 3. *Pour tout groupe de Lie connexe G il existe un groupe de Lie simplement connexe $\tilde{G}$ localement isomorphe et dépendant fonctoriellement de G.*

Le groupe $\tilde{G}$ est défini de façon unique à un isomorphisme près.

Deux groupes de Lie connexes G et H sont localement isomorphes si et seulement si les groupes $\tilde{G}$ et $\tilde{H}$ sont isomorphes.

Un groupe de Lie connexe est localement isomorphe au groupe G si et seulement s'il est isomorphe au groupe quotient $\tilde{G}/K$ du groupe $\tilde{G}$ par un sous-groupe discret (donc central) invariant K.

Démonstration. Les deux premières affirmations sont, en vertu de la proposition 3, une autre formulation du théorème 2.

Si $P \to G$ et $P \to H$ sont des isomorphismes locaux et $\tilde{P} \to P$ un revêtement universel simplement connexe, les applications composées $\tilde{P} \to G$ et $\tilde{P} \to H$ seront aussi des revêtements universels. Donc $\tilde{P} \approx \tilde{G} \approx \tilde{H}$. Réciproquement, si $\tilde{G} \approx \tilde{H}$, on a les isomorphismes locaux $P \to G$ et $P \to H$ avec $P = \tilde{G}$. Ceci prouve la troisième affirmation.

La quatrième affirmation résulte de la troisième en vertu de la proposition 3. □

Du théorème 3 il s'ensuit, en particulier, que *la relation d'isomorphisme local est une relation d'équivalence.*

Le groupe $\tilde{G}$ s'appelle *groupe de revêtement simplement connexe* du groupe G.

LEÇON 10

Isomorphismes locaux et isomorphismes de localisations.— Théorème de Cartan.— Diagramme final des catégories et des foncteurs.— Réduction du théorème de Cartan.— Globalisabilité des groupuscules plongeables.— Réduction du théorème de Cartan au théorème d'Ado.

Le théorème 3 de la leçon précédente nous donne une description assez satisfaisante des classes des groupes de Lie localement isomorphes, mais présente le défaut de décrire des classes autres que celles qui ont été introduites dans la leçon 3, c'est-à-dire des classes de groupes de Lie qui ne sont pas isomorphes dans la catégorie GR-LOC des groupuscules. Nous devons donc prouver accessoirement la coïncidence de ces deux notions d'isomorphisme local. La clef nous est fournie par la proposition suivante:

Proposition 1. *Soient G et H des groupes de Lie connexes, U un voisinage connexe de l'unité dans le groupe G et $\varphi: U \to H$ une application différentiable telle que $\varphi(xy) = \varphi(x)\varphi(y)$ pour tous éléments $x, y \in U$ pour lesquels $xy \in U$. Sous ces conditions, si le groupe G est simplement connexe, il existe un seul homomorphisme $\Phi: G \to H$ prolongeant l'application φ, c'est-à-dire tel que*

$$\Phi|_U = \varphi.$$

Pour prouver cette proposition, considérons le sous-ensemble D, du produit $G \times G$, de tous les couples $(x_1, x_2) \in G \times G$ tels que $x_1x_2^{-1} \in U$. Ce sous-ensemble est visiblement ouvert et contient la diagonale $\Delta \subset G \times G$. Comme D est la réunion d'ensembles connexes de la forme $\{x\} \times Ux$, $x \in G$, coupant chacun la diagonale Δ qui est aussi un ensemble connexe (homéomorphe au groupe G), il s'ensuit que *l'ensemble D est connexe.*

On dit qu'un sous-ensemble V d'un groupe G est *petit* si $V \times V \subset$ $\subset D$, c'est-à-dire si $VV^{-1} \subset U$, et qu'un sous-ensemble W du produit $G \times H$ des groupes G et H est *distingué* si pour tout point

$(x, y) \in W$ il existe dans G un petit voisinage V du point x tel que

$$(v, \varphi(vx^{-1})\, y) \in W \quad \text{pour tout point } v \in V.$$

Il est clair que

a) un ensemble vide est distingué;

b) le produit $G \times H$ est distingué;

c) la réunion de toute famille d'ensembles distingués est distinguée;

d) l'intersection de toute famille finie d'ensembles distingués est distinguée.

On peut donc traiter les ensembles distingués comme des ensembles ouverts d'une nouvelle topologie sur le produit $G \times H$. On désignera par $\tilde{X}$ le produit $G \times H$ muni de cette topologie.

Soit

$$\pi: \ \tilde{X} \to G$$

une application définie par la formule

$$\pi(x, y) = x, \qquad (x, y) \in \tilde{X}.$$

Il est évident que l'ensemble $\pi^{-1}(V)$ est distingué pour tout ouvert $V \subset G$ et que l'ensemble $\pi(W)$ est ouvert pour tout ensemble distingué $W \subset X$. Ceci exprime que l'application π est continue et ouverte.

Soient V un petit ensemble ouvert de G, x_0 un élément de V et y_0 un élément de H. Désignons par $W(x_0, V, y_0)$ l'ensemble de tous les couples de la forme $(x, \varphi(xx_0^{-1})\, y_0)$, où $x \in V$. Il est immédiat de voir que l'ensemble $W(x_0, V, y_0)$ est distingué. En effet, si $(x, y) \in W(x_0, V, y_0)$, c'est-à-dire si $x \in V$ et $y = \varphi(xx_0^{-1})\, y_0$, alors pour tout point $v \in V$, on a l'égalité

$$\varphi(vx^{-1})\, y = \varphi(vx^{-1})\, \varphi(xx_0^{-1})\, y_0 = \varphi(vx_0^{-1})\, y_0,$$

qui montre que $(v, \varphi(vx^{-1})\, y) \in W(x_0, V, y_0)$.

Il est évident que $(x_0, y_0) \in W(x_0, V, y_0)$, c'est-à-dire que $W(x_0, V, y_0)$ est un voisinage du point (x_0, y_0) dans l'espace $\tilde{X}$. Il est manifeste que ce voisinage est homéomorphe au voisinage V par l'application π. Par ailleurs, il est immédiat de voir que l'ensemble $\pi^{-1}(V)$ est la réunion de tous les ensembles de la forme $W(x_0, V, y)$, où $x_0 \in W$ est fixe, et $y \in H$ parcourt le groupe H tout entier, ces ensembles étant deux à deux disjoints (s'il existe un point $(x, y) \in W(x_0, V, y_1) \cap W(x_0, V, y_2)$, alors $\varphi(xx_0^{-1})\, y_1 = y = \varphi(xx_0^{-1})\, y_2$ et, par suite, $y_1 = y_2$).

Ceci signifie que tout petit ensemble ouvert $V \subset G$ est revêtu sans pli par l'application π. Donc, tout élément du groupe G possède manifestement un petit voisinage, ce qui prouve que *l'application $\pi: \tilde{X} \to G$ est un faible revêtement.* Donc, d'après le lemme 2 de la

leçon 8, pour toute composante $\widetilde{X}_0$ de l'espace $\widetilde{X}$, *l'application*

$$\pi_0 = \pi\,|_{\widetilde{X}_0} : \widetilde{X}_0 \to G$$

est un revêtement. Prenons pour composante $\widetilde{X}_0$ celle qui contient le point (e_G, e_H), où e_G et e_H sont les unités des groupes G et H respectivement.

Souvenons-nous maintenant que le groupe G est par hypothèse simplement connexe. Donc, chacun de ses revêtements est trivial, c'est-à-dire est un homéomorphisme. En particulier, le revêtement π_0 est un homéomorphisme. L'homéomorphisme réciproque π_0^{-1} envoie tout point $x \in G$ en un point de l'espace $\widetilde{X}_0$, de la forme (x, y), où $y \in H$. Donc, en posant $\Phi(x) = y$, on obtient une application continue définie de façon unique

$$\Phi\colon G \to H.$$

Donc, pour tout point $x \in G$, le point $\Phi(x) \in H$ est caractérisé de façon unique par le fait que $(x, \Phi(x)) \in \widetilde{X}_0$. D'où, en particulier, $\Phi(e_G) = e_H$.

Soit D^* le sous-ensemble d'un ensemble D, formé des points $(x_1, x_2) \in D$ tels que

$$(1) \qquad \Phi(x_2) = \varphi(x_2x_1^{-1})\,\Phi(x_1).$$

Il est évident que $\Delta \subset D^*$, donc D^* n'est pas vide. Par ailleurs, il est aisé de voir que *pour tout petit ensemble ouvert connexe* $V \subset G$, *on a l'inclusion*

$$V \times V \subset D^*.$$

En effet, pour tout point $x_0 \in V$, l'ensemble distingué $W(x_0, V, \Phi(x_0))$ est connexe (car homéomorphe à V) et contient le point $(x_0, \Phi(x_0)) \in \widetilde{X}_0$. Donc, $W(x_0, V, \Phi(x_0)) \subset \widetilde{X}_0$, c'est-à-dire que pour tout point $x \in V$, le point $(x, \varphi(xx_0^{-1})\,\Phi(x_0))$ est situé dans $\widetilde{X}_0$. Mais, d'après ce qui précède, tout point de $\widetilde{X}_0$ peut être représenté d'une seule façon sous la forme $(x, \Phi(x))$. Par conséquent, $\Phi(x) = \varphi(xx_0^{-1})\,\Phi(x_0)$, ce qui équivaut à l'inclusion $(x_0, x) \in D^*$. Donc, $V \times V \subset D^*$. □

Il est aisé de voir, par ailleurs, que *si le produit $V \times V'$ de deux petits ouverts connexes V et V' rencontre D^*, il est contenu dans D^**. En effet, si $(x, x_0) \in D^*$, $(x_0, x_0') \in D^*$ et $(x_0', x') \in D^*$, alors $(x, x') \in D^*$ (car $\Phi(x') = \varphi(x'x_0'^{-1})\ \Phi(x_0') = \varphi(xx_0'^{-1}) \times \varphi(x_0'x_0^{-1}) \times \Phi(x_0) = \varphi(x'x_0^{-1})\,\varphi(x_0x^{-1})\,\Phi(x) = \varphi(x'x^{-1})\,\Phi(x)$). Par ailleurs, si $(x_0, x_0') \in V \times V'$ et $(x, x') \in V \times V'$, alors d'après ce qui a été prouvé, $(x, x_0) \in D^*$ et $(x_0', x') \in D^*$. Donc, si de plus $(x_0, x_0') \in D^*$, alors $(x, x') \in D^*$. □

Les ensembles $V \times V'$ formant de toute évidence une base de D, on en déduit immédiatement que l'ensemble D^* est ouvert et fermé dans D. Comme D est connexe et D^* non vide, ceci n'est possible que si $D^* = D$. Donc, *l'égalité* (1) *est valable pour tout point* $(x_1, x_2) \in D$.

Tous les éléments sont réunis pour prouver la proposition 1.

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 1. L'unicité de l'homomorphisme Φ résulte immédiatement du fait (cf. lemme 3 de la leçon 9) que le groupe G est engendré par le voisinage U. Il nous faut donc en prouver seulement l'existence.

Montrons que cet homomorphisme n'est autre que l'application Φ construite ci-dessus.

En faisant $x_1 = e_G$ et $x_2 = x$ dans l'égalité (1), on obtient aussitôt $\Phi(x) = \varphi(x)$ pour tout élément $x \in U$. Donc, on prouvera la proposition 1 lorsqu'on aura montré que l'application Φ est un homomorphisme, c'est-à-dire que pour tous éléments $x, x' \in G$ on aura

$$\Phi(x, x') = \Phi(x) \cdot \Phi(x'). \tag{2}$$

Remarquons à cet effet que si $x \in U$, l'égalité (2) est valable pour tout $x' \in G$ (car $(x', xx') \in D$ et, par suite, $\Phi(xx') = \varphi(x)\,\Phi(x') = \Phi(x)\,\Phi(x')$). Par ailleurs, le groupe G étant connexe, il est engendré par le voisinage $U \cap U^{-1}$ de l'unité. donc tout élément $x \in G$ peut être représenté sous la forme

$$x = x_1 x_2 \ldots x_n,$$

où $x_1, x_2, \ldots, x_n \in U$. Donc, par une récurrence évidente, on établit que

$$\Phi(xx') = \Phi(x_1) \ldots \Phi(x_n)\,\Phi(x') \tag{3}$$

pour tout élément $x' \in G$. En particulier, pour $x' = e_G$,

$$\Phi(x) = \Phi(x_1) \ldots \Phi(x_n),$$

ce qui combiné à (3) prouve (2). □

L'application φ de la proposition 1 n'est autre qu'un homomorphisme du groupuscule G_{loc} dans le groupuscule H_{loc}, où G_{loc} et H_{loc} sont les images des groupes G et H par le foncteur de localisation

$$\text{GR}_0\text{-DIFF} \to \text{GR-LOC}, \tag{4}$$

et l'application Φ, un homomorphisme du groupe G dans le groupe H transformé par le foncteur (4) en l'homomorphisme φ. Donc, la proposition 3 n'affirme rien d'autre que — G étant simplement connexe — la condition d'univalence complète du foncteur (4) est remplie pour les groupes G et H. Si donc l'on se borne à la sous-

catégorie complète GR_{00}-DIFF de la catégorie GR_0-DIFF composée des groupes de Lie simplement connexes, cette condition sera remplie sans aucune réserve. Donc, *le foncteur de localisation*

(5) $$GR_{00}\text{-DIFF} \to GR\text{-LOC}$$

est complètement univalent sur la catégorie GR_{00}-DIFF.

Mais il est immédiat de voir que tout foncteur univalent établit une correspondance biunivoque entre les isomorphismes et donc, en particulier, envoie uniquement des objets isomorphes dans des objets isomorphes. Appliqué au foncteur (5) cela signifie que *les groupes de Lie simplement connexes sont isomorphes si et seulement si leurs localisations le sont.*

Ceci nous permet de répondre par l'affirmative à la question posée au début de la leçon : *des groupes de Lie connexes sont localement isomorphes au sens de la leçon* 3 (*c'est-à-dire possèdent des localisations isomorphes*) *si et seulement s'ils sont localement isomorphes au sens du théorème* 3 *de la leçon* 9. En effet, tout isomorphisme local au sens de la définition 3 de la leçon 9 sera visiblement un isomorphisme de localisations. Donc, en particulier, les localisations de tout groupe de Lie G et de son groupe de revêtement universel $\widetilde{G}$ sont isomorphes :

$$G_{\text{loc}} \approx \widetilde{G}_{\text{loc}}.$$

Par conséquent, si $G_{\text{loc}} \approx H_{\text{loc}}$, alors $\widetilde{G}_{\text{loc}} \approx \widetilde{H}_{\text{loc}}$, et, par suite, $\widetilde{G} \approx \widetilde{H}$ d'après ce qui précède, de sorte que les groupes G et H seront localement isomorphes. ☐

Donc, *dans le théorème* 3 *de la leçon* 9, *par « isomorphisme local » des groupes de Lie, on peut comprendre l'isomorphisme de leurs localisations.*

Cela signifie que ce théorème répond entièrement à la question de l'inversibilité du foncteur de localisation (4) :

$$G_{\text{loc}} \approx H_{\text{loc}} \quad \textit{si et seulement si} \quad \widetilde{G} \approx \widetilde{H}.$$

En particulier, *le foncteur* (5) *est inversible* (à un isomorphisme près) *sur la catégorie* GR_{00}-DIFF *des groupes simplement connexes.*

Mais ce résultat n'est pas encore complet, car on ne sait pas si tout groupuscule de Lie est la localisation d'un groupe de Lie, c'est-à-dire si le foncteur (5) établit l'équivalence des catégories.

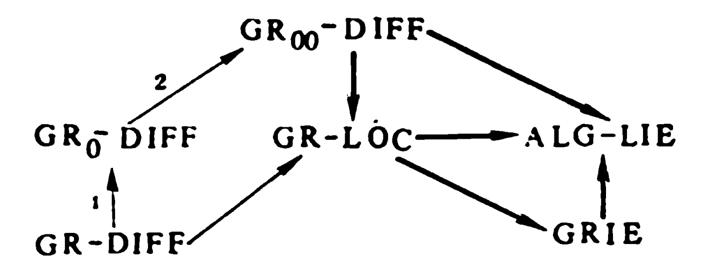
La réponse à cette question est affirmative :

Théorème 1 (Cartan). *Le foncteur* (5) *établit l'équivalence de la catégorie* GR_{00}-DIFF *des groupes de Lie connexes simplement connexes et de la catégorie* GR-LOC *des groupuscules de Lie.*

Corollaire. *La catégorie des groupes de Lie connexes simplement connexes est équivalente à la catégorie des algèbres de Lie de dimension finie sur le corps* $\mathbb{R}$. *Cette équivalence est réalisée par un foncteur de Lie.*

Ce corollaire est le point culminant de la théorie développée. Il permet en effet de ramener tout problème relatif aux groupes de Lie connexes et simplement connexes à un problème homologue sur les algèbres de Lie, problème qui en principe est bien plus simple, car étant l'analogue « linéaire » du problème primitif.

Les catégories et foncteurs considérés plus haut forment le diagramme commutatif



dans lequel les flèches épaisses désignent les foncteurs réalisant l'équivalence des catégories.

Le chiffre 1 représente le foncteur de passage à la composante de l'unité. Ce foncteur envoie dans un même groupe tous les groupes qui sont les extensions du groupe de Lie connexe donné au moyen d'un groupe discret quelconque.

Le chiffre 2 représente le foncteur de passage à un groupe de revêtement universel. Ce foncteur envoie dans un même groupe tous les groupes qui sont les groupes quotients du groupe de Lie simplement connexe donné par les sous-groupes invariants (centraux) discrets.

Ce diagramme contient toute l'information nécessaire sur les relations entre les groupes et les algèbres de Lie.

Pour achever notre théorie il nous reste donc à prouver seulement le théorème de Cartan. Mais avec ce théorème la situation est assez particulière et même pas satisfaisante du tout.

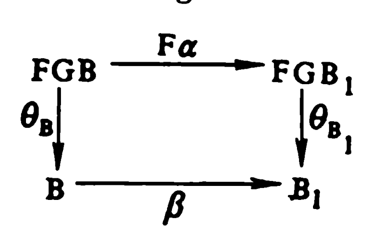
La voie la plus naturelle pour prouver le théorème de Cartan consiste à construire un foncteur de la catégorie des groupuscules (ou, ce qui est équivalent, de la catégorie des algèbres de Lie) dans la catégorie des groupes de Lie simplement connexes, qui soit quasiréciproque du foncteur de localisation. Mais, à ce jour, en dépit des tentatives, qui, il faut le croire, sont nombreuses (en tout cas les mathématiciens de Moscou se sont sérieusement penchés sur le sujet), aucune construction foncièrement « naturelle » (s'appuyant

sur les seules notions fondamentales) de ce foncteur n'a pu être proposée. J.-P. Serre va même jusqu'à penser qu'elle n'existe pas (cf. [9]; à noter à propos que Serre appelle le théorème de Cartan « troisième théorème de Lie »). Toutes les démonstrations connues (en fait, elles sont au nombre de deux) du théorème de Cartan ne revêtent pas un caractère fonctoriel (= « naturel ») et soulèvent une réprobation « *in petto* ». Il faut espérer que le dernier mot n'a pas encore été dit.

Ces démonstrations reposent sur le lemme catégoriel suivant:

Lemme 1. *Soient* A *et* B *des catégories arbitraires, et soit* $F: \mathsf{A} \to \mathsf{B}$ *un foncteur complètement univalent de la catégorie* A *dans la catégorie* B. *Si pour tout objet* B *de la catégorie* B, *il existe un objet* A *de la catégorie* A *tel que l'objet* FA *soit isomorphe à l'objet* B, *le foncteur* F *est quasiinversible* (*réalise l'équivalence des catégories*).

D é m o n s t r a t i o n. Pour tout objet B de la catégorie B, choisissons et figeons l'objet A de la catégorie A mentionné dans le lemme et désignons-le par GB. Figeons aussi un isomorphisme $\theta_B: FGB \to B$. Le foncteur F étant complètement univalent, pour tout morphisme $\beta: B \to B_1$, il existe un morphisme $\alpha: GB \to GB_1$ et un seul tel que l'on ait le diagramme commutatif



On pose $\alpha = G\beta$.

De l'unicité du morphisme α il s'ensuit aussitôt que les correspondances construites $\mapsto B\ GB$ et $\beta \mapsto G\beta$ forment un foncteur G: $\mathsf{B} \to \mathsf{A}$, les isomorphismes θ_B établissant visiblement un isomorphisme entre le foncteur FG et le foncteur identique Id_{B}. Par ailleurs, pour tout objet A de la catégorie A, l'égalité $F(\alpha_A) = \theta_{FA}$ détermine un morphisme $\alpha_A: GFA \to A$ qui est un isomorphisme et de plus, ce qui est aisé à voir, les isomorphismes α_A forment un isomorphisme entre le foncteur GF et le foncteur Id.

Donc, le foncteur G est quasiréciproque du foncteur F. □

En vertu de ce lemme général, pour prouver le théorème de Cartan, il suffit de construire pour tout groupuscule de Lie au moins un groupe de Lie pour lequel ce groupuscule (ou un groupuscule isomorphe) soit un voisinage de l'unité. De plus (et ceci est un gage de succès), on peut ignorer la fonctorialité et laisser la porte ouverte à tout arbitraire dans la construction.

Définition 1. On dira qu'un groupuscule de Lie est *globalisable* s'il est isomorphe à la localisation d'un groupe de Lie (c'est-à-dire est isomorphe à un voisinage de l'unité de ce groupe).

Donc, pour prouver le théorème de Cartan, il nous faut simplement montrer que *tout groupuscule de Lie est globalisable.*

Définition 2. On dit qu'un groupuscule de Lie K est *plongeable* s'il est sous-groupuscule d'un groupe de Lie G, plus exactement, de la localisation G_{loc} de ce groupe.

La première étape de la démonstration du théorème de Cartan est la proposition suivante:

Proposition 2. *Tout groupuscule de Lie K plongeable est globalisable.*

Commençons par démontrer une proposition analogue pour les groupes topologiques.

Soient G un groupe topologique connexe et K un sous-espace de G contenant l'unité e de G et tel que $xy \in K$ et $x^{-1} \in K$ pour tous éléments x, $y \in K \cap U_0$, où U_0 est un voisinage de e.

Lemme 2. *Il existe un groupe topologique H et un homomorphisme injectif $i: H \to G$ qui applique homéomorphiquement un voisinage V_0 de l'unité du groupe H sur un voisinage de l'unité e dans K, c'est-à-dire sur un ensemble de la forme $K \cap U_{00}$, où U_{00} est un voisinage de l'unité e dans le groupe G.*

D é m o n s t r a t i o n. On dira qu'un sous-ensemble $A \subset G$ est *séparé* de e si dans G il existe un voisinage U de e tel que $A \cap U = \varnothing$. Il est clair que *si des ensembles A et B sont séparés de e, il en est de même de l'ensemble $A \cup B$.*

Pour tout élément $g \in G$, on désignera par A^g l'ensemble $g^{-1}Ag$ de tous les éléments de la forme $g^{-1}ag$, $a \in A$. Il est évident que *l'ensemble A^g est séparé de e si et seulement si l'ensemble A est séparé de e.*

Soit H l'ensemble de tous les éléments $g \in G$ pour lesquels la différence symétrique

$$K^g \triangle K = (K^g \setminus K) \cup (K \setminus K^g) = (K^g \cup K) \setminus (K^g \cap K)$$

est séparée de e. Cet ensemble est un sous-groupe du groupe G (ou, plus exactement, du groupe abstrait correspondant G_{abstr}), puisque $K^{g^{-1}} \triangle K = (K^g \triangle K)^{g^{-1}}$ et $K^{g_1 g_2} \triangle K \subset (K^{g_1} \triangle K)^{g_2} \cup (K^{g_2} \triangle K)$.

Soit U_{00} un voisinage de l'unité e du groupe G, tel que $xy \in U_0$ pour tous éléments x, $y \in U_{00}$. Alors $g^{-1}xg \in K$ et $gxg^{-1} \in K$ pour tous éléments x, $g \in K \cap U_{00}$. Donc, si $g \in K \cap U_{00}$, $y \in K^g \cap U_{00}^g$, et, par suite, $x = gyg^{-1} \in K \cap U_{00}$, alors $y = g^{-1}xg \in K$. Par conséquent, $K^g \cap U_{00}^g \subset K \cap U_{00}$ et donc $K^g \cap V_0 \subset K \cap V_0$, où $V_0 = U_{00} \cap U_{00}^g$. Réciproquement, si $x \in K \cap U_{00}$, alors $y = gxg^{-1} \in K$ et, par suite, $x = g^{-1}yg \in K^g$. Par conséquent,

$K \cap U_{00} \subset K^g \cap U_{00}$ et $K \cap V_0 \subset K^g \cap V_0$. Donc, $K \cap V_0 = K^g \cap V_0$, c'est-à-dire que $(K^g \triangle K) \cap V_0 = \varnothing$. Ceci prouve que *pour tout élément* $g \in K \cap U_{00}$, *l'ensemble* $K^g \triangle K$ *est séparé de* e, c'est-à-dire que $g \in H$. Donc, $K \cap U_{00} \subset H$.

Munissons H d'une topologie en prenant pour voisinages de l'unité tous les ensembles de la forme $V = K \cap U$, où $U \subset U_{00}$ est un voisinage arbitraire de l'unité de G, contenu dans U_{00}. Une vérification immédiate montre que H est muni *ipso facto* d'une topologie qui en fait un groupe topologique. Sous ces conditions l'injection $i\colon H \to G$ sera un homomorphisme injectif de groupes topologiques et cette application sera un homéomorphisme sur le voisinage $V_0 = K \cap U_{00}$.

Ceci achève la démonstration du lemme 1. □

A noter qu'en général H ne sera pas un sous-groupe du groupe G, c'est-à-dire que l'application $i\colon H \to G$ ne sera pas un homéomorphisme sur $i(H)$. Si par exemple $K = \{e\}$, alors H est un groupe de G pour la topologie discrète.

Nous pouvons maintenant prouver la proposition 2.

Démonstration de la proposition 2. Supposons que le groupuscule K est un sous-groupuscule du groupe de Lie G, donc (cf. leçon 7), dans G il existe une carte $(U_0, h) = (U_0, x^1, \ldots, x^n)$ telle que $K \cap U_0$ est défini par les équations

$$x^{m+1} = 0, \ldots, x^n = 0.$$

En traitant G comme un groupe topologique, on peut appliquer le lemme 2 à K. D'après ce lemme, il existe un groupe topologique H et un homomorphisme continu injectif $i\colon H \to G$ qui applique homéomorphiquement un voisinage V_0 de l'unité du groupe H sur l'ensemble $K \cap U_{00}$, où U_{00} est un voisinage de l'unité du groupe G. Ceci étant, on peut, sans nuire à la généralité, admettre que $U_{00} = U_0$.

Dans ces conditions, le couple $(V_0, k) = (V_0, y^1, \ldots, y^m)$, où $y^j(v) = x^j(iv)$, $j = 1, \ldots, n$, sera de toute évidence une carte de H contenant l'unité $e \in V$ et, par suite, pour tout élément $a \in H$, le couple $(aV_0, k \circ L_{a^{-1}})$ sera une carte de H contenant a. Si $aV_0 \cap bV_0 \neq \varnothing$, alors sur $(k \circ L_{a^{-1}})(aV_0 \cap bV_0) \subset \mathbb{R}^m$ l'application $(k \circ L_{b^{-1}}) \circ (k \circ L_{a^{-1}})^{-1} = k \circ L_{b^{-1}a} \circ k^{-1}$ sera (puisque $\mathbb{R}^m \subset \mathbb{R}^n$) la restriction de l'application différentiable $h \circ L_{i(b^{-1}a)} \circ h^{-1}$ et donc sera elle-même une application différentiable. Ceci prouve que les cartes de la forme $(aV_0, k \circ L_{a^{-1}})$ sont compatibles et forment ainsi un atlas de H. Une vérification immédiate montre que H est un groupe de Lie pour la structure différentiable définie par cet atlas, et l'application i, un homomorphisme différentiable qui est un

difféomorphisme du voisinage V_0 sur $K \cap U_0$. Donc, $H_{loc} \approx K$ et le groupuscule K est globalisable. □

Soit maintenant K un groupuscule de Lie et soit $\mathfrak{k} = \mathfrak{l}(K)$ son algèbre de Lie.

Dans la leçon 19 on démontrera le théorème suivant:

Théorème d'Ado. *Toute algèbre de Lie de dimension finie (sur un corps* $\mathbb{K}$ *de caractéristique* 0) *est isomorphe à une algèbre de Lie de matrices.*

D'après ce théorème, on peut, sans nuire à la généralité, admettre que $\mathfrak{k}$ est une algèbre de Lie de matrices et, par suite, que le groupuscule correspondant $E\mathfrak{k} \approx K$ est composé de matrices, c'est-à-dire est un sous-groupuscule du groupe de Lie GL(n). Donc, le groupuscule K est plongeable et, par suite, globalisable.

Ce raisonnement ramène le théorème de Cartan à celui d'Ado. Nous prouverons le théorème d'Ado (donc celui de Cartan) dans les leçons 17 à 20. Pour l'instant penchons-nous sur les sous-groupes des groupes de Lie.

LEÇON 11

Sous-variétés de variétés différentiables.— Sous-groupes de groupes de Lie.— Variétés intégrales de sous-fibrés intégrables.— Variétés intégrales maximales.— Principe de la démonstration du théorème 1.— Structure locale des sous-variétés.— Unicité de la structure d'une sous-variété localement redressable à base dénombrable.— Sous-variétés de variétés à base dénombrable.— Les groupes de Lie connexes possèdent une base dénombrable.— Redressabilité locale des variétés intégrales maximales.— Démonstration du théorème 1.

Le tore à deux dimensions T^2 peut être représenté comme le groupe quotient du groupe additif $\mathbb{R}^2$ par son lattis $\mathbb{Z}^2$ des points à coordonnées entières. Donc, toute droite de $\mathbb{R}^2$ passant par le point (0, 0) nous donnera un sous-groupe de T^2. Si le coefficient directeur de cette droite de $\mathbb{R}^2$ est rationnel (par exemple égal à m/n), l'image de cette droite dans T^2 sera un cercle parcourant le tore m fois le long d'un méridien et n fois le long d'un parallèle. Mais si ce coefficient est irrationnel, le sous-groupe correspondant de T^2 (qui est appelé *hélice irrationnelle du tore*) sera partout dense dans T^2, et tout voisinage de l'un quelconque de ses points contiendra, pour la topologie induite, un voisinage qui est la somme d'un nombre dénombrable d'intervalles fermés disjoints. Ce sous-groupe ne peut donc être une variété.

Cet exemple explique pourquoi nous adoptons la définition suivante qui apparemment semble trop générale:

Définition 1. Une variété différentiable N est une *sous-variété* d'une variété différentiable M si:

a) tout point de N appartient à M, de sorte qu'est définie l'injection ι: $N \rightarrow M$;

b) l'application ι: $N \rightarrow M$ est différentiable (et donc, en particulier, continue);

c) la différentielle $(d\iota)_a \colon \mathsf{T}_a(N) \to \mathsf{T}_a(M)$ de l'application ι en $a \in N$ est un monomorphisme.

En général tout vecteur $A \in \mathsf{T}_a(N)$ est identifié au vecteur $(d\iota)_a A$, c'est-à-dire qu'on admet que $\mathsf{T}_a(N)$ est un sous-espace de l'espace $\mathsf{T}_a(M)$.

Il est évident que toute sous-variété ouverte est une sous-variété au sens de la définition 1. Une telle sous-variété est un sous-espace, c'est-à-dire que l'application ι est un homéomorphisme sur son image. Mais signalons que ceci n'est pas le cas pour une hélice irrationnelle du tore.

Définition 2. On dit qu'un groupe de Lie H est un *sous-groupe* d'un groupe de Lie G si la variété H_{diff} est sous-variété de la variété G_{diff} et le groupe H_{abstr}, sous-groupe du groupe G_{abstr}. On dit qu'un sous-groupe H d'un groupe de Lie G est un *sous-groupe invariant* si le sous-groupe H_{abstr} l'est.

En particulier, les hélices irrationnelles du tore sont ses sous-groupes.

Soient $\mathfrak{g} = \mathfrak{l}(G)$ et $\mathfrak{h} = \mathfrak{l}(H)$ les algèbres de Lie des groupes de Lie G et H. Si H est un sous-groupe du groupe de Lie G, l'homomorphisme $\mathfrak{l}(\iota) \colon \mathfrak{h} \to \mathfrak{g}$ identifié à l'application $d(\iota)_e \colon \mathsf{T}_e(H) \to \to \mathsf{T}_e(G)$ est un monomorphisme. En général, l'algèbre de Lie $\mathfrak{h}$ est identifiée à son image dans $\mathfrak{g}$ par ce monomorphisme. Donc, en vertu de cette identification, *les algèbres de Lie des sous-groupes du groupe de Lie G sont sous-algèbres de l'algèbre de Lie $\mathfrak{g}$ du groupe G.*

La correspondance $H \mapsto \mathfrak{h}$ entre les sous-groupes $H \subset G$ et les sous-algèbres $\mathfrak{h} \subset \mathfrak{g}$ sera appelée *correspondance de Lie.* On n'étudiera que les sous-groupes c o n n e x e s H, puisque les algèbres de Lie d'un groupe et de sa composante de l'unité sont confondues. Il s'avère que la correspondance de Lie est biunivoque dans ces conditions. On a donc le théorème suivant:

Théorème 1. *La correspondance de Lie $H \mapsto \mathfrak{h}$ est une correspondance biunivoque entre les sous-groupes connexes du groupe de Lie G et les sous-algèbres de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$.*

On peut évidemment admettre que le groupe G est connexe dans ce théorème.

Nous prouverons le théorème 1 par étapes.

Soient M une variété différentiable et E un sous-fibré intégrable du fibré tangent $\mathsf{T}(M)$ (cf. leçon 7).

Définition 3. On dit qu'une sous-variété N de la variété M est une *variété intégrale* du sous-fibré E si l'injection $\iota \colon N \to M$

est intégrale par rapport à E, c'est-à-dire si

$$E_a = \mathsf{T}_a(N)$$

pour tout $a \in N$.

Il est immédiat de voir que *le fibré E est intégrable (au sens de la définition* 10 *de la leçon* 7) *si et seulement si par tout point a de la variété M il passe au moins une variété intégrale de ce fibré.* En effet, il est évident que cette condition est une condition suffisante d'intégrabilité. Réciproquement, supposons que le fibré E est intégrable. Il est alors complètement intégrable (corollaire de la proposition 6 de la leçon 7), c'est-à-dire que (définition 13 de la leçon 7) la variété M possède un atlas composé de cartes $(U, x^1, \ldots, x^n)$ telles que pour tout point $a \in U$, les vecteurs $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots, \left(\frac{\partial}{\partial x^m}\right)_a$ forment une base dans l'espace E_a. Pour tout point $\xi = (\xi^1, \ldots \ldots, \xi^{n-m}) \in \mathbb{R}^{n-m}$ désignons par V_ξ l'ensemble des points $a \in U$ dont les coordonnées satisfont les conditions

$$x^{m+1} = \xi^1, \quad \ldots, \quad x^n = \xi^{n-m}.$$

Cet ensemble (quand il n'est pas vide) est muni de façon naturelle d'une structure de variété différentiable difféomorphe à un ouvert de $\mathbb{R}^m$. Il est évident que cette variété est une sous-variété dans U (donc dans M) et de plus en tout point $a \in V_\xi$, le sous-espace $\mathsf{T}_a(V_\xi)$ est engendré par les vecteurs $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots, \left(\frac{\partial}{\partial x^m}\right)_a$. Donc, $\mathsf{T}_a(V_\xi) = E_a$, de sorte que V_ξ est une variété intégrale du sous-fibré E. Pour achever la démonstration, il reste à remarquer que pour tout point $a \in U$, la sous-variété V_ξ, où $\xi = (x^{m+1}(a), \ldots \ldots, x^n(a))$, contient a. □

Dans la suite on aura assez souvent affaire à des sous-variétés de la forme V_ξ. On les appellera *sous-variétés plates* de la carte U.

On dira que des sous-variétés N_1 et N_2 de la variété M, ayant un point $a \in M$ en commun sont *confondues localement en* a s'il existe une sous-variété N_0, ouverte aussi bien dans N_1 que dans N_2, telle que $a \in N_0$.

Il est immédiat de voir que *deux sous-variétés intégrales N_1 et N_2 d'un sous-fibré E intégrable, passant par a, sont confondues localement en a.* En effet, soient $\iota_1 : N_1 \to M$ et $\iota_2 : N_2 \to M$ des injections. D'après le lemme 2 de la leçon 7, il existe des voisinages V_1 et V_2 du point a dans les sous-variétés N_1 et N_2 et un difféomorphisme $\beta : V_2 \to V_1$, tels que $\iota_2 = \iota_1 \circ \beta$ sur V_1. Mais, ι_1 et ι_2 étant les restrictions de l'application identique $M \to M$, l'égalité $\iota_2 = \iota_1 \circ \beta$ n'est possible que si $\beta = \mathrm{id}$ (donc si $V_1 = V_2$). Ce qui achève la démonstration, puisque l'égalité $\iota_1 = \iota_2$ dans un voisinage du point a exprime justement que les variétés N_1 et N_2 sont localement confondues en a. □

Considérons l'ensemble $\mathfrak{N}$ de toutes les variétés intégrales d'un fibré vectoriel intégrable E. D'après ce qui précède, si deux sous-variétés N_1, $N_2 \in \mathfrak{N}$ se coupent, alors pour tout point $a \in N_1 \cap N_2$, il existe une sous-variété N_0, ouverte aussi bien dans N_1 que dans N_2. telle que $a \in N_0 \subset N_1 \cap N_2$. Comme $N_0 \in \mathfrak{N}$ (une sous-variété ouverte d'une variété de $\mathfrak{N}$ est visiblement contenue dans $\mathfrak{N}$), ceci montre que l'on peut prendre la famille $\mathfrak{N}$ pour base d'une nouvelle topologie sur M (les ensembles ouverts de cette topologie seront des réunions de sous-variétés de $\mathfrak{N}$). On désignera par M_E une variété M munie de cette topologie.

Il est immédiat de voir que *l'application identique* $M_E \to M$ *est continue*, c'est-à-dire que tout ensemble ouvert $U \subset M$ est ouvert aussi dans M_E. En effet, soit $a \in U$, et soit $a \in N \in \mathfrak{N}$. L'ensemble N étant localement connexe, la composante N_0 de l'ensemble $N \cap U$, qui contient le point a, est ouverte dans N, c'est-à-dire est une sous-variété ouverte de la variété N. Donc, $N_0 \in \mathfrak{N}$. Ce qui exprime que U est la réunion de variétés de $\mathfrak{N}$, donc qu'il est ouvert dans M_E. □

D'autre part, il est aisé de voir que la topologie de l'espace M_E induit sur toute variété $N \in \mathfrak{N}$ (qui, par définition, est un sous-ensemble ouvert de l'espace M_E) la topologie primitive de la variété N, de sorte que, en d'autres termes, *toute variété intégrale* $N \in \mathfrak{N}$ *est un sous-espace ouvert de l'espace* M_E. En effet, tout voisinage du point $a \in N$ pour la topologie de la variété N est une variété intégrale de $\mathfrak{N}$ et, par suite, un ensemble ouvert dans M_E. Réciproquement, tout voisinage U de a pour la topologie induite par celle de l'espace M_E contient une variété intégrale $N_1 \in \mathfrak{N}$ telle que $a \in N_1$. Les variétés N et N_1 étant localement confondues en a, il existe une variété $N_0 \in \mathfrak{N}$ qui contient le point a et qui est une sous-variété ouverte aussi bien de la variété N que de la variété N_1. La variété N_0 sera précisément un voisinage de a pour la topologie de la variété N, contenu dans le voisinage U. Donc, la topologie primitive de la variété N est confondue avec la topologie induite sur N par celle de l'espace M_E. □

Soit maintenant W une composante connexe de l'espace M_E (munie de la topologie induite). Toute variété intégrale connexe N contenant un point $a \in W$ est contenue dans W (car N est connexe aussi en tant que sous-espace de M_E).

En particulier, W contient chaque carte U du point a dans la variété N. U sera un voisinage de a dans W aussi, puisque U est ouvert dans M_E en tant que variété intégrale. Ceci permet de considérer chaque carte (U, h) en a de la variété N comme une carte de la variété W. On obtient ainsi sur W une structure différentiable compatible avec la topologie.

La variété différentiable W construite est visiblement une variété intégrale connexe du sous-fibré E, qui contient le point a

et qui possède la propriété suivante: toute variété intégrale connexe du sous-fibré E renfermant le point a est contenue et ouverte dans W.

Définition 4. On dit qu'une variété intégrale d'un sous-fibré E est *maximale* si elle est connexe et n'est contenue dans aucune autre variété intégrale connexe de ce sous-fibré.

On voit, en particulier, que la variété intégrale construite W est maximale.

On a ainsi prouvé la proposition suivante:

Proposition 1. *Par tout point $a \in M$ il passe une seule variété intégrale maximale W du sous-fibré intégrable E. Toute variété intégrale connexe du sous-fibré E passant par un point a est une sous-variété ouverte de la variété W.* □

Revenons aux groupes de Lie et considérons une sous-algèbre $\mathfrak{h}$ de l'algèbre de Lie $\mathfrak{g}$ d'un groupe de Lie connexe G. On rappelle que par $\mathfrak{a}(G)$ on désigne l'algèbre de Lie de dimension infinie de tous les champs de vecteurs sur G. Cette algèbre est un module sur l'algèbre $\mathcal{F}(G)$ de toutes les fonctions différentiables sur G. Pour tout sous-fibré $E \subset \mathsf{T}(G)$, les champs de vecteurs situés dans E forment un sous-module $\mathfrak{a}(E)$ du module $\mathfrak{a}(G)$. L'algèbre de Lie $\mathfrak{g}$, donc $\mathfrak{l}$, sont des sous-algèbres de $\mathfrak{a}(G)$. Donc, le sous-module $\mathcal{F}(G)\mathfrak{h}$ engendré par la sous-algèbre $\mathfrak{h}$ sera une sous-algèbre de $\mathfrak{a}(G)$. De plus $\mathcal{F}(G)\mathfrak{g} = \mathfrak{a}(G)$.

La première étape de la démonstration du théorème 1 consistera à prouver que *pour toute sous-algèbre $\mathfrak{h}$ de l'algèbre de Lie $\mathfrak{g}$. il existe un sous-fibré $E^{\mathfrak{h}}$ du fibré tangent $\mathsf{T}(G)$, tel que*

$$\mathfrak{a}(E^{\mathfrak{h}}) = \mathcal{F}(G)\mathfrak{h}. \tag{1}$$

En effet, soit $E^{\mathfrak{h}}_a$ le sous-espace de $\mathsf{T}_a(G)$ des vecteurs de la forme X_a, où $X \in \mathfrak{h}$ (l'algèbre de Lie $\mathfrak{g}$ sera interprétée ici comme l'algèbre des champs de vecteurs invariants à gauche; si $\mathfrak{g} = \mathsf{T}_e(G)$, le sous-espace $E^{\mathfrak{h}}_a$ est l'image de l'espace $\mathfrak{h} \subset \mathsf{T}_e(G)$ par l'application $(dL_a)_e$). Une vérification immédiate montre que la réunion

$$E^{\mathfrak{h}} = \bigcup_{a \in G} E^{\mathfrak{h}}_a$$

de ces sous-espaces est un sous-fibré du fibré $\mathsf{T}(G)$ possédant la propriété (1). □

On dira que le sous-fibré $E^{\mathfrak{h}}$ est *engendré* par la sous-algèbre $\mathfrak{h}$.

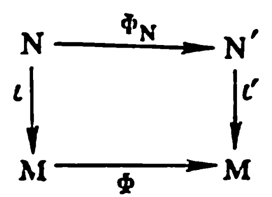
D'après ce qui précède, de (1) il s'ensuit que $\mathfrak{a}(E^{\mathfrak{h}})$ est une sous-

algèbre de l'algèbre $\mathfrak{a}(G)$, c'est-à-dire que le sous-fibré $E^{\mathfrak{h}}$ est involutif. Donc, en vertu du théorème de Frobenius (proposition 5 de la leçon 7), *le sous-fibré* $E^{\mathfrak{h}}$ *est intégrable.*

Si la sous-agèbre $\mathfrak{h}$ est l'algèbre de Lie d'un sous-groupe connexe H, alors H sera visiblement une variété intégrale du sous-fibré $E^{\mathfrak{h}}$ passant par le point e (nous verrons plus bas que cette variété est maximale). Donc, pour toute sous-algèbre $\mathfrak{h}$, il est naturel d'introduire la variété intégrale maximale H du sous-fibré $E^{\mathfrak{h}}$ qui passe par e et d'essayer de prouver qu'elle est un sous-groupe de Lie dont l'algèbre de Lie est $\mathfrak{h}$.

Pour cela il faut d'abord prouver que H_{abstr} *est un sous-groupe du groupe* G_{abstr}, c'est-à-dire que $xy \in H$ et $x^{-1} \in H$ quels que soient $x, y \in H$. Les considérations générales suivantes nous permettent d'établir ceci sans peine.

Soient $\Phi: M \to M'$ un difféomorphisme, N une sous-variété de la variété M. Considérons l'image N' de la sous-variété N par l'application Φ. Etant donné que Φ est une application bijective de N sur N', on peut l'utiliser pour transporter la structure différentiable de N à N'. On fera ainsi de N' une variété différentiable et l'application Φ induira un difféomorphisme $\Phi_N: N \to N'$ bouclant le diagramme commutatif



dont les flèches verticales figurent des injections. La commutativité de ce diagramme exprime que $\iota' = \Phi \circ \iota \circ \Phi_N^{-1}$, d'où il s'ensuit immédiatement que l'application ι' est différentiable et sa différentielle $(d\iota')_{a'}$ en tout point $a' \in N'$ est un monomorphisme. Ceci signifie que N' *est une sous-variété de la variété* M. Cette sous-variété sera appelée *image* de la sous-variété N par le difféomorphisme Φ.

Si maintenant E est un sous-fibré intégrable quelconque du fibré $\mathsf{T}(M)$, son image E' par le difféomorphisme $\mathsf{T}(\Phi): \mathsf{T}(M) \to \mathsf{T}(M')$ sera un sous-fibré intégrable du fibré $\mathsf{T}(M')$, et *l'image de toute variété intégrale maximale du sous-fibré E par le difféomorphisme* Φ *sera une variété intégrale maximale du sous-fibré* E'.

Comme le sous-fibré E' est confondu visiblement avec $E = E^{\mathfrak{h}}$ pour $M = M' = G$, $E = E^{\mathfrak{h}}$ et $\Phi = L_x$ (ou $\Phi = I$, où $I: G \to G$ est l'application $x \mapsto x^{-1}$), on en déduit que l'image par Φ de la sous-variété intégrale maximale H du sous-fibré $E^{\mathfrak{h}}$ qui passe par e est la sous-variété intégrale maximale qui passe par le point Φe, c'est-

à-dire que cette image est H si $\Phi e \in H$. Comme, de toute évidence, la condition $\Phi e \in H$ est remplie pour $\Phi = I$, ainsi que pour $\Phi = L_x$, $x \in H$, ceci prouve que $I(H) = H$ et $L_x(H) = H$ pour $x \in H$, c'est-à-dire que *H est un sous-groupe*.

Pour achever la démonstration, il reste à montrer que H est un groupe de Lie, c'est-à-dire que les applications $(x, y) \mapsto xy$ et $x \mapsto x^{-1}$ sont différentiables. La différentiabilité de l'application $x \mapsto x^{-1}$ résulte de ce qui précède. S'agissant de l'application $(x, y) \mapsto xy$, tout ce qu'on peut dire pour l'instant c'est qu'elle est différentiable seulement par rapport à x ou à y séparément.

Contre toute attente, la démonstration de la différentiabilité de l'application $(x, y) \mapsto xy$ par rapport à x et y est un problème assez compliqué qui implique des considérations topologiques fines et un grand travail préparatoire. Ceci nous contraint une fois de plus à revenir aux bases de la théorie.

Définition 5. On appelle *rang* d'une application différentiable $\Phi: N \to M$ d'une variété N dans une variété M en un point $a \in N$ le rang de sa différentielle

$$(d\Phi)_a: \mathsf{T}_a(N) \to \mathsf{T}_a(M)$$

considérée comme une application linéaire de l'espace vectoriel $\mathsf{T}_a(N)$ dans l'espace vectoriel $\mathsf{T}_a(M)$.

Soient $y^1, \ldots, y^m$ des coordonnées locales sur la variété N au point a, et $x^1, \ldots, x^n$ des coordonnées locales sur la variété M au point Φa. Soient par ailleurs

$$\varphi^1(y), \ldots, \varphi^n(y), \qquad y = (y^1, \ldots, y^n),$$

les fonctions définissant l'application Φ dans ces coordonnées. Comme la matrice de l'application linéaire $(d\Phi)_a$ dans les bases $\left(\frac{\partial}{\partial y^1}\right)_a, \ldots, \left(\frac{\partial}{\partial y^m}\right)_a$ et $\left(\frac{\partial}{\partial x^1}\right)_{\Phi a}, \ldots, \left(\frac{\partial}{\partial x^n}\right)_{\Phi a}$ est la matrice jacobienne dont les éléments sont les dérivées partielles

$$(2) \qquad \left(\frac{\partial \varphi^i}{\partial y^j}\right)_a, \quad i = 1, \ldots, n \quad j = 1, \ldots, m,$$

calculées au point $(y^1(a), \ldots, y^m(a)) \in \mathbb{R}^m$, *le rang de l'application Φ en a est égal à celui de la matrice d'éléments* (2).

Par continuité, il s'ensuit de là que le rang r' de l'application Φ en un point quelconque a' d'un voisinage U du point a est supérieur à son rang r en a:

$$r' \geqslant r.$$

Si $r' = r$ pour tout point $a' \in U$, l'application Φ s'appelle *localement plane* en a.

Quitte à renuméroter les coordonnées, on peut admettre que

(3) $$\det\left|\left(\frac{\partial\varphi^i}{\partial y^j}\right)_a\right| \neq 0 \text{ pour } i,\ j=1,\ \ldots,\ r.$$

Considérons les fonctions $y'^1, \ldots, y'^m$ définies au voisinage du point a par les formules

$$y'^j(y^1, \ldots, y^m) = \begin{cases} \varphi^j(y^1, \ldots, y^m) & \text{pour } 1 \leqslant j \leqslant r, \\ y^j & \text{pour } r+1 \leqslant j \leqslant m. \end{cases}$$

Le jacobien $\frac{D(y'^1, \ldots, y'^m)}{D(y', \ldots, y^m)}$ étant non nul au point a en vertu de la condition (3), les fonctions $y'^1, \ldots, y'^m$ sont des coordonnées locales dans un voisinage du point a. Dans ces coordonnées, l'application Φ est définie par les formules

$$x^i = \begin{cases} y'^i & \text{pour } 1 \leqslant i \leqslant r, \\ \varphi'^i(y'^1, \ldots, y'^m) & \text{pour } r+1 \leqslant i \leqslant n. \end{cases}$$

Si l'application Φ est localement plane au point a, il est aisé de voir que dans ces formules les fonctions $\varphi'^i(y'^1, \ldots, y'^m)$, $r+1 \leqslant i \leqslant n$, ne dépendent pas des coordonnées $y'^{r+1}, \ldots, y'^m$, c'est-à-dire sont de la forme

$$\varphi'^{r+1}(y'^1, \ldots, y'^r), \ldots, \varphi'^n(y'^1, \ldots, y'^r).$$

Donc, les formules

$$x'^i = \begin{cases} x^i & \text{pour } 1 \leqslant i \leqslant r, \\ x^i - \varphi'^i(x^1, \ldots, x^r) & \text{pour } r+1 \leqslant i \leqslant n, \end{cases}$$

définissent dans un voisinage du point Φa des coordonnées locales $x'^1, \ldots, x'^n$ telles que l'application Φ est définie dans les coordonnées $y'^1, \ldots, y'^n$ et $x'^1, \ldots, x'^n$ par les formules

$$x'^i = \begin{cases} y'^i & \text{pour } 1 \leqslant i \leqslant r, \\ 0 & \text{pour } r+1 \leqslant i \leqslant n. \end{cases}$$

Nous avons ainsi prouvé la proposition suivante:

Proposition 2. *Si une application différentiable $\Phi\colon N \to M$ est localement plane en un point $a \in N$ et son rang est égal à r en ce point, alors il existe sur les variétés N et M des coordonnées locales $y^1, \ldots, y^m$ et $x^1, \ldots, x^n$ (définies respectivement dans des voisinages des points a et Φa) telles que*

$$x^i \circ \Phi = \begin{cases} y^i & \textit{pour } 1 \leqslant i \leqslant r, \\ 0 & \textit{pour } r+1 \leqslant i \leqslant n. \end{cases} \quad \square$$

Appliquée à l'injection $\iota: N \to M$ d'une sous-variété N dans M (qui est de toute évidence une application localement plane), cette proposition affirme que *pour tout point* $a \in N$, *il existe sur la variété* M *une carte* $(U, x^1, \ldots, x^n)$, $a \in U$, *telle que les fonctions*

$$x^1 \circ \iota, \ldots, x^m \circ \iota \tag{4}$$

sont des coordonnées locales dans un voisinage V *du point* a *sur* N, *et les fonctions*

$$x^{m+1} \circ \iota, \ldots, x^m \circ \iota$$

identiquement nulles (dans V*).*

La dernière assertion signifie que *la sous-variété* N *est localement confondue en* a *avec la sous-variété plate* $V_0 = V$ *de la carte* U.

Remarque 1. L'exemple de l'hélice irrationnelle du tore montre que d'une façon générale $V_0 \neq U \cap N$.

Comme l'injection $\iota: N \to M$ est différentiable, pour toute fonction f différentiable sur M, sa restriction $f \circ \iota$ à N est différentiable sur N (si, bien sûr, cette restriction existe, c'est-à-dire si l'ensemble $W(f)$ de définition de f coupe N). Réciproquement, considérons une fonction g différentiable sur N. Supposons que l'ensemble $W(g)$ de définition de cette fonction contient la carte V du point a sur la variété N, construite ci-dessus. Alors

$$g = \hat{g}(x^1 \circ \iota, \ldots, x^m \circ \iota) \text{ sur } V,$$

où $\hat{g}$ est une fonction différentiable de m variables. Sur une carte U du point a sur la variété M, définissons une fonction f au moyen de la formule

$$f = \hat{g}(x^1, \ldots, x^m).$$

Il est évident que f est une fonction différentiable et $f \circ \iota = g$. Nous avons ainsi établi le lemme suivant:

Lemme 1. *Tout point d'une sous-variété* N *possède un voisinage* V *tel que toute fonction différentiable sur* N *est la restriction à* V *d'une fonction différentiable sur* M. □

Corollaire. *Si un sous-ensemble* N *d'une variété différentiable* M *est muni d'une structure de sous-variété, alors cette structure est unique pour la topologie de* N, *c'est-à-dire que toute autre structure de sous-variété sur* N *induira une autre topologie sur* N.

Démonstration. D'après le lemme 1, toute carte différentiable $(W, y^1, \ldots, y^m)$ de N est composée d'un ouvert W et de fonctions $y^1, \ldots, y^m$ ayant un jacobien non nul et qui sont localement des restrictions à N de fonctions différentiables sur M.

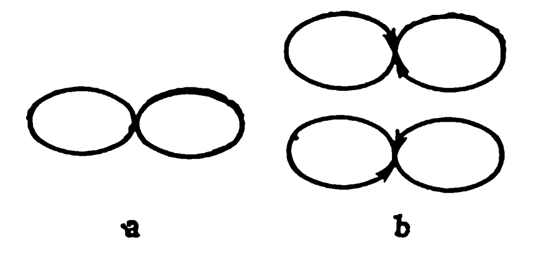
Donc, deux structures quelconques de sous-variétés sur N, induisant une même topologie sur N, auront des cartes identiques et, par suite, seront confondues. □

En faisant varier la topologie, on peut certes obtenir des structures différentes de sous-variété sur N. Par exemple, on peut munir tout sous-ensemble $N \subset M$ de la topologie discrète et en faire *ipso facto* une sous-variété de dimension zéro.

On obtient un exemple moins trivial en considérant le carré ouvert de $\mathbb{R}^2$: $-1 < x < 1$, $-1 < y < 1$. C'est une sous-variété ouverte à deux dimensions pour la topologie induite. Mais on peut en faire une sous-variété à une dimension en le munissant d'une topologie dont les ouverts sont des ensembles dont l'intersection avec chaque intervalle vertical $x = x_0$, $-1 < y < 1$ est un ensemble ouvert (dans cet intervalle). La sous-variété obtenue est composée d'un nombre non dénombrable de composantes difféomorphes chacune à la droite $\mathbb{R}$. Cette sous-variété sera appelée *carré débité*.

De telles pathologies seront exclues si l'on impose à la sous-variété d'admettre une base dénombrable.

Il existe cependant d'autres possibilités de faire varier la topologie. Considérons par exemple dans le plan $\mathbb{R}^2$ l'ensemble (en forme de huit) de la figure *a*.



Ce sous-ensemble n'est pas une sous-variété pour la topologie induite par la topologie du plan. On peut toutefois en faire une sous-variété, et ce de deux manières, en le munissant de topologies plus fortes représentées formellement sur la figure *b*. Dans les deux cas, la sous-variété obtenue est difféomorphe à la droite $\mathbb{R}$ et, par suite, à l'inverse des sous-variétés précédentes, est connexe et possède une base dénombrable.

Définition 6. Une sous-variété N (de dimension m) d'une variété M (de dimension n) est dite *localement redressable* en un point $a \in N$ s'il existe une carte (U, h) de la variété M et un sous-ensemble $\Xi \subset \mathbb{R}^{n-m}$ tels que $a \in U$ et l'intersection $U \cap N$ est la réunion, pour la topologie induite par celle de la sous-variété N, de sous-variétés plates (cf. page 182) V_ξ, $\xi \in \Xi$, de la carte (U, h):

$$U \cap N = \bigcup_{\xi \in \Xi} V_\xi .$$

La sous-variété N sera dite *localement redressable* si elle l'est en chacun de ses points.

Les sous-variétés de la figure b ne sont pas localement redressables.

L'ensemble « paramétrant » Ξ peut, en général, être quelconque, et le fait de son existence ne nous couvre pas encore contre les pathologies. Par exemple, le carré débité est localement redressable, et pour chacun de ses points l'ensemble Ξ est un intervalle fermé. Mais cet ensemble ne peut être trop grand dans des situations moins pathologiques. Par exemple, du fait que toutes les sous-variétés V_ξ sont visiblement ouvertes dans N, il s'ensuit immédiatement que *si une sous-variété localement redressable N possède une base dénombrable, alors pour chacun de ses points a l'ensemble Ξ est au plus dénombrable.*

Par ailleurs, de la définition de la topologie induite, il s'ensuit immédiatement que *si la topologie de la sous-variété N est induite par la topologie de la variété enveloppante M, alors la variété N est localement redressable et pour chacun de ses points l'ensemble Ξ peut être un singleton.*

Donc, les sous-variétés localement redressables peuvent être traitées comme les extensions de sous-variétés munies d'une topologie induite.

Pour éviter les mentions, on conviendra une fois pour toutes que les cartes (U, h) sont *cubiques*, c'est-à-dire que l'ensemble $h(U) \subset \subset \mathbb{R}^n$ est un cube

$$|x^1| < c, \ldots, |x^n| < c.$$

Le nombre c sera appelé *semi-largeur de la carte* (U, h).

Sous ces conditions, toute sous-variété plate V_ξ sera connexe (et non vide) pour $|\xi^1| < c, \ldots, |\xi^{n-m}| < c$.

Donc, en particulier, pour une sous-variété N localement redressable en a, toutes les variétés plates V_ξ, $\xi \in \Xi$, qui forment l'intersection $U \cap N$, seront les composantes connexes de cette intersection pour la topologie induite par celle de la sous-variété N. L'exemple du carré débité montre qu'en général ceci est mis en défaut pour la topologie de la variété enveloppante M. Mais, *si la sous-variété N possède une base dénombrable, les sous-variétés V_ξ seront les composantes de l'intersection $U \cap N$ également pour la topologie induite par celle de la variété M.* En effet, l'application $U \cap N \to \Xi$ définie par la formule

$$a \mapsto (x^{m+1}(a), \ldots, x^n(a)), \tag{2}$$

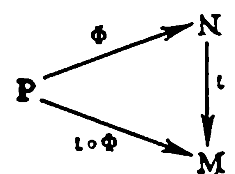
est visiblement continue (pour la topologie de la variété M). Donc, elle envoie composante de l'ensemble $U \cap N$ dans composante de l'ensemble Ξ. Mais les composantes de Ξ sont des points, puisque celui-ci est dénombrable. Donc, chaque composante de l'ensemble

$U \cap N$ est contenue dans l'image réciproque V_ξ d'un point $\xi \in \Xi$ par l'application (2), donc est confondue avec V_ξ. □

Lemme 2. *Soient P et M des variétés différentiables, N une sous-variété localement redressable de la variété M possédant une base dénombrable, $\iota: N \to M$ l'injection de N dans M et $\Phi: P \to N$ une application de P dans N telle que la composée $\iota \circ \Phi: P \to M$ soit une application continue. Alors l'application Φ est aussi continue.*

Si, par ailleurs, l'application $\iota \circ \Phi$ est différentiable, il en est de même de l'application Φ.

Cette situation est illustrée par le diagramme



D é m o n s t r a t i o n. Soient $b \in P$ et $a = \Phi(b)$.

Par définition, il existe dans la variété M une carte (U, h) (pour laquelle $h(a) = 0$) telle que l'intersection $U \cap N$ est la réunion, pour la topologie induite, d'un nombre dénombrable de sous-variétés plates connexes V_ξ, $\xi \in \Xi$. L'application $\iota \circ \Phi$ étant continue, il existe dans la variété P un voisinage W du point b dont l'image $(\iota \circ \Phi)(W)$ est contenue dans U et, par suite, dans $U \cap N$. Ceci étant, on peut sans restreindre la généralité admettre que le voisinage W est connexe. Sous ces conditions (une application continue envoie ensemble connexe dans ensemble connexe), l'image $(\iota \circ \Phi)(W)$ du voisinage W par l'application $\iota \circ \Phi$ sera aussi connexe et donc contenue dans une composante V_ξ de l'ensemble $U \cap N$. Comme $\Phi(b) = a$ et $a \in V_0$, cette composante est nécessairement $V_0 = V$. Donc, pour le voisinage V du point a dans la variété N il existe un voisinage W du point b dans la variété P, tel que $\Phi(W) \subset V$. Les voisinages V formant un système fondamental (une base) de voisinages du point a dans la variété N, ceci prouve que l'application $\Phi: P \to N$ est continue en b. Comme le point $b \in P$ est arbitraire, l'application Φ est continue.

Supposons maintenant que l'application $\iota \circ \Phi$ est di férentiable. Pour prouver que l'application Φ est différentiable, il faut montrer que pour toute fonction g différentiable sur N, la fonction $g \circ \Phi$ est différentiable sur P. Mais, en vertu du lemme 1, la fonction g a pour expression locale $f \circ \iota$, où f est une fonction différentiable sur M. Donc, la fonction $g \circ \Phi$ est confondue localement avec la fonction différentiable $f \circ (\iota \circ \Phi)$ et, par suite, est différentiable. □

L'exemple des variétés de la figure b montre que le lemme 2 est faux sans la condition de redressement local.

Nous pouvons maintenant démontrer le théorème d'unicité annoncé:

Proposition 3. *Si un sous-ensemble N d'une variété différentiable peut être muni d'une structure de sous-variété localement redressable admettant une base dénombrable, il ne peut l'être que d'une seule manière.*

D é m o n s t r a t i o n. Soient N' et N'' des sous-variétés localement redressables admettant une base dénombrable, dont l'ensemble sous-jacent est N. L'application identique $N' \to N''$ est alors justiciable du lemme 2. Donc, cette application est différentiable. L'application identique $N'' \to N'$ est différentiable pour des raisons analogues. Donc, les structures différentiables de N' et N'' sont confondues. □

Corollaire. *Soient M une variété différentiable, N une sous-variété de M localement redressable admettant une base dénombrable et $\Phi: M \to M$ un difféomorphisme de M sur M tel que $\Phi a \in N$ et $\Phi^{-1}a \in N$ pour tout point $a \in N$. Sous ces conditions, l'application $N \to N$ induite par le difféomorphisme Φ est aussi un difféomorphisme.*

D é m o n s t r a t i o n. Soit N' l'image de la sous-variété N par le difféomorphisme Φ. Il est évident que N' est aussi une sous-variété localement redressable admettant une base dénombrable. De plus, elle est confondue par hypothèse avec N en tant qu'ensemble. Donc, d'après la proposition 2, elle sera confondue avec N et en tant que variété. Par conséquent, le difféomorphisme $N \to N'$ induit par le difféomorphisme Φ sera bien un difféomorphisme $N \to N$. □

Les résultats acquis appellent la question suivante: sous quelles conditions une sous-variété admet-elle une base dénombrable? On se propose de montrer que pour qu'une sous-variété connexe possède une base dénombrable, il suffit que la variété ambiante M en possède une.

Lemme 3. *Supposons que pour un espace topologique connexe X il existe un recouvrement ouvert $\{U_\alpha\}$ tel que:*

1) *tout ensemble U_α (muni de la topologie induite) est un espace admettant une base dénombrable;*

2) *pour tout α_0 il existe un nombre au plus dénombrable d'ensembles U_α coupant U_{α_0}.*

Alors, l'espace X possède une base dénombrable.

D é m o n s t r a t i o n. Il suffit de prouver que X est la réunion d'une sous-famille dénombrable (ou finie) du recouvrement $\{U_\alpha\}$. Figeons un $U_{\alpha_0} \neq \varnothing$ et considérons les éléments U_α de ce recouvre-

ment pour lesquels existe une suite finie

$$U_{\alpha_1}, U_{\alpha_2}, \ldots, U_{\alpha_n}$$

d'éléments du recouvrement $\{U_\alpha\}$ tels que $U_{\alpha_n} = U_\alpha$ et $U_{\alpha_{i-1}} \cap \cap U_{\alpha_i} \neq \varnothing$ pour tout $i = 1, \ldots, n$. De la condition 2) il s'ensuit immédiatement que ces éléments forment une sous-famille au plus dénombrable du recouvrement $\{U_\alpha\}$. Soit X' leur réunion. Il est manifeste que X' est ouvert (et pas vide). Il est aussi fermé, puisque si $x \in \overline{X}'$ et $x \in U_\alpha$, alors $X' \cap U_\alpha \neq \varnothing$, et, par suite, U_α coupe un élément de la sous-famille construite, donc appartient lui-même à cette sous-famille. Par conséquent $U_\alpha \subset X$ et $x \in X$. Etant un sous-ensemble non vide ouvert et fermé de l'espace connexe X, l'ensemble X' est confondu avec X tout entier. Donc, X possède une base dénombrable. □

Lemme 4. *Supposons que pour un espace connexe X, il existe un recouvrement ouvert dénombrable $\{U_k\}$ dont chaque élément U_k est la réunion d'ensembles ouverts disjoints U_{k,α_n} possédant une base dénombrable. Alors, l'espace X possède aussi une base dénombrable.*

Démonstration. Il suffit de prouver que le recouvrement ouvert $\{U_{k,\alpha_k}\}$ satisfait la condition 2) du lemme 3. Le recouvrement $\{U_k\}$ étant dénombrable, il suffit de montrer que quels que soient k, l et α_k, il existe un nombre au plus dénombrable d'ensembles U_{l,α_l} coupant U_{k,α_k}. Supposons le contraire, c'est-à-dire supposons qu'il existe un famille non dénombrable d'ensembles U_{l,α_l} coupant U_{k,α_k}. En prenant un point dans chaque intersection, on obtient dans U_{k,α_k} un sous-ensemble non dénombrable de points isolés (on rappelle que les ensembles U_{l,α_l} sont disjoints par hypothèse). Comme tout sous-ensemble discret d'un espace admettant une base dénombrable est au plus dénombrable, l'existence d'un tel sous-ensemble non dénombrable contredit le fait que U_{k,α_k} est un espace admettant une base dénombrable. Donc, U_{k,α_k} est coupé par un nombre au plus dénombrable d'ensembles U_{l,α_l}. □

Corollaire 1. *Tout espace revêtant un espace admettant une base dénombrable admet lui-même une base dénombrable.*

Démonstration. Les images réciproques des éléments d'un recouvrement dénombrable composé d'ensembles revêtus sans pli forment un recouvrement de l'espace de revêtement, satisfaisant les conditions du lemme 4. □

Corollaire 2. *Si pour un espace connexe X, il existe un homéomorphisme local*

$$f\colon X \to \mathbb{R}^m,$$

alors X possède une base dénombrable.

13—01642

D é m o n s t r a t i o n. Soit $\{U_k\}$ une base dénombrable de l'espace $\mathbb{R}^m$ formée d'ouverts connexes (par exemple de parallélépipèdes), et soient $V_{k,\alpha}$ des sous-ensembles ouverts de l'espace X appliqués au moyen de f sur U_k (α parcourt ici un ensemble d'indices A, dépendant de k et susceptible d'être vide pour certains k). Dire que f est un homéomorphisme local revient à dire que les ensembles

$$V_k = \bigcup_{\alpha \in A_k} V_{k,\alpha}$$

recouvrent X tout entier. Pour pouvoir appliquer le lemme 4, il nous faut donc montrer seulement que les ensembles connexes ouverts $V_{k,\alpha}$ sont les composantes des ensembles V_k, c'est-à-dire que quels que soient α, $\beta \in A_k$, si $V_{k,\alpha} \cap V_{k,\beta} \neq \emptyset$, alors $V_{k,\alpha} = V_{k,\beta}$.

Soit $W = V_{k,\alpha} \cap V_{k,\beta} \neq \emptyset$, et soient

$$g_\alpha: U_k \to V_{k,\alpha}, \quad g_\beta: U_k \to V_{k,\beta}$$

les homéomorphismes réciproques de $f|_{V_{k,\alpha}}$ et $f|_{V_{k,\beta}}$. Si $x_n \in W$ et $x = \lim x_n$, alors $f(x) = \lim f(x_n)$ et $g_\beta(f(x_n)) = x_n$. Donc,

$$g_\beta(f(x)) = \lim g_\beta(f(x_n)) = \lim x_n = x.$$

Donc, si $x \in V_{k,\alpha}$ (et, par suite, $f(x) \in U_k$), alors $x \in V_{k,\beta}$, c'est-à-dire que $x \in W$. Ceci exprime que W est fermé dans $V_{k,\alpha}$. Comme W est, de plus, ouvert dans $V_{k,\alpha}$ (et pas vide), et que $V_{k,\alpha}$ est connexe (car homéomorphe à l'ensemble connexe U_k), ceci n'est possible que pour $V_{k,\alpha} = W$. On montre de façon analogue que $V_{k,\beta} = W$. Donc, $V_{k,\alpha} = V_{k,\beta}$. □

Corollaire 3. *Toute sous-variété connexe N de la variété $\mathbb{R}^n$ possède une base dénombrable.*

D é m o n s t r a t i o n. Soient $x^1, \ldots, x^n$ des coordonnées dans $\mathbb{R}^n$ (dans une certaine base), et soit ι l'injection $N \to \mathbb{R}^n$. Considérons pour un m-uple $\alpha = (i_1, \ldots, i_m)$ d'indices $1, \ldots, n$ (m étant la dimension de la variété N) un sous-ensemble V_α de la variété N composé de points $a \in N$ au voisinage de chacun desquels les fonctions $x^{i_1} \circ \iota, \ldots, x^{i_m} \circ \iota$ sont des coordonnées locales. Il est évident que V_α est ouvert dans N. De plus (cf. proposition 3), tout point $a \in N$ appartient au moins à un ensemble V_α. Soit V'_α une composante de l'ensemble V_α. L'application $V'_\alpha \to \mathbb{R}^m$ définie par la formule

$$a \mapsto (x^{i_1}(a), \ldots, x^{i_m}(a)), \quad a \in V'_\alpha,$$

est un homéomorphisme local par construction. Donc, d'après le corollaire 2, l'ensemble connexe V'_α possède une base dénombrable. Par conséquent, le recouvrement ouvert fini $\{V_\alpha\}$ de la variété N satisfait les conditions du lemme 4. Par suite, la variété N possède une base dénombrable. □

Nous pouvons maintenant prouver le théorème annoncé plus haut.

Proposition 4. *Toute sous-variété connexe N d'une variété M admettant une base dénombrable est une variété qui admet une base dénombrable.*

D é m o n s t r a t i o n. Soit $\{U_k\}$ un recouvrement dénombrable de la variété M, formé de cartes, et soit $V_k = U_k \cap N$. Les ensembles V_k sont ouverts dans N et tout point $a \in N$ appartient au moins à l'un d'eux. Chaque composante V'_k de V_k est une variété différentiable connexe, difféomorphe à une sous-variété de l'espace $\mathbb{R}^m$. Donc, d'après le corollaire 3, cette composante possède une base dénombrable. Ceci exprime que le recouvrement ouvert $\{V_k\}$ de la variété N satisfait toutes les conditions du lemme 4. Donc, la variété N possède une base dénombrable. □

L'applicabilité des résultats acquis ci-dessus aux groupes de Lie est assurée par la proposition suivante:

Proposition 5. *Tout groupe de Lie connexe G possède une base dénombrable.*

D é m o n s t r a t i o n. Soit U un voisinage de l'unité du groupe, difféomorphe à un ouvert de $\mathbb{R}^n$ et tel que $U^{-1} = U$. Choisissons dans U un ensemble Y dénombrable partout dense et considérons l'ensemble Z de tous les éléments du groupe G de la forme $y_1 y_2 \ldots$ $\ldots y_k$, où $y_1, \ldots, y_k \in Y$ (k est arbitraire). Il est manifeste que l'ensemble Z est dénombrable. Le groupe G étant connexe, il est engendré par le voisinage U, c'est-à-dire que tout élément x de G est de la forme $x_1 x_2 \ldots x_k$, où $x_1, x_2, \ldots, x_k \in U$. Soit $x_{(i)} =$ $= x_1 x_2 \ldots x_i$, $i = 1, \ldots, k$ (en particulier, $x_{(1)} = x_1$ et $x_{(k)} =$ $= x$). De la continuité de la multiplication de G, on déduit par une récurrence évidente que l'unité de G possède un voisinage V tel que

$$V \cdot x_{(1)} V x_{(1)}^{-1} \cdot x_{(2)} V x_{(2)}^{-1} \cdot \ldots \cdot x_{(k)} V x_{(k)}^{-1} \subset U.$$

L'ensemble Y étant partout dense dans U, le voisinage Vx_i du point $x_i \in U$ contient un point $y_i \in Y$. Soit $v_i = y_i x_i^{-1} \in V$. Alors

$$y_1 y_2 \ldots y_k = v_1 x_1 \cdot v_2 x_2 \cdot \ldots \cdot v_k x_k =$$
$$= v_1 \cdot x_{(1)} v_2 x_{(1)}^{-1} \cdot x_{(1)} v_3 x_{(2)}^{-1} \cdot \ldots \cdot x_{(k-1)} v_k x_{(k)}^{-1} \cdot x_{(k)} = ux,$$

où $u \in U$. Ceci exprime qu'un point $z = y_1 \ldots y_k \in Z$ est tel que $x \in Uz$ (on rappelle que $U^{-1} = U$ par hypothèse). Ceci prouve que les ouverts de la forme Uz, $z \in Z$, recouvrent G. Comme Z est dénombrable et les ensembles Uz (qui sont homéomorphes à U) possèdent une base dénombrable, on en déduit que G possède aussi une base dénombrable. □

Corollaire. *Toute sous-variété connexe d'un groupe de Lie connexe est une variété admettant une base dénombrable.*

En particulier, une variété intégrale maximale H du sous-fibré $E\mathfrak{h}$ est une sous-variété possédant une base dénombrable. Ceci étant, il est immédiat de voir qu'elle est localement redressable. En effet, *dans toute variété M, chaque variété intégrale maximale W d'un sous-fibré intégrable E est localement redressable*, puisque la variété M est recouvrable par des cartes telles que chacune de leurs sous-variétés plates V_ξ est une variété intégrale du sous-fibré E ; c'est pourquoi l'intersection de W avec une telle carte sera, puisque W est maximale, la réunion de certaines de ces sous-variétés, ce qui exprime par définition que la variété W est localement redressable. □

Remarque 2. Que la variété intégrale H soit localement redressable résulte encore du lemme général suivant:

Lemme 5. *Une sous-variété connexe H d'un groupe de Lie G pour lequel l'ensemble H_{abstr} est un sous-groupe du groupe G_{abstr} est localement redressable.*

D é m o n s t r a t i o n. On sait que dans G il existe une carte $(U, x^1, \ldots, x^n)$ au point e telle que la sous-variété plate

$$V: x^{m+1} = 0, \quad \ldots, \quad x^n = 0$$

est un voisinage du point e dans la sous-variété H et un sous-groupuscule du groupuscule U. Dans l'algèbre de Lie $\mathfrak{l}(G) = \mathfrak{l}(U)$, à ce sous-groupuscule est associée la sous-algèbre $\mathfrak{h} = \mathfrak{l}(V)$.

Sans nuire à la généralité, on peut visiblement supposer que les coordonnées $x^1, \ldots, x^n$ sont des coordonnées canoniques définies par une décomposition de la forme $\mathfrak{l}(G) = \mathfrak{h} \oplus \mathfrak{k}$. Alors, les sous-variétés plates V_ξ de la carte U seront (cf. leçon 7) les classes $aV = aH \cap U$ du sous-groupuscule V. L'intersection $U \cap H$ étant la réunion des classes aV pour lesquelles $a \in U \cap H$, ceci prouve que $U \cap H$ est la réunion de certaines sous-variétés plates V_ξ. Donc, la sous-variété H est localement redressable au point e.

Soit maintenant a un élément quelconque de H. Considérons le difféomorphisme $L_a: G \to G$. Ce difféomorphisme applique H sur lui-même et envoie le point e en a. Donc, la sous-variété H est localement redressable en a aussi. □

La sous-variété H admettant une base dénombrable et étant localement redressable, il vient, en vertu du lemme 2, que *toute application $\Phi: P \to H$ d'une variété différentiable quelconque P dans la variété intégrale H, telle que la composée*

$$\iota \circ \Phi: P \to G$$

de Φ et de l'injection $\iota: H \to G$ soit différentiable, est différentiable.

Appliquons cette affirmation à la variété $P = H \times H$ et à l'application $\mu_H: (x, y) \mapsto xy$. Comme on le diagramme commutatif

$$\begin{array}{ccc} H\times H & \xrightarrow{\mu_H} & H \\ {\scriptstyle\iota\times\iota}\downarrow & & \downarrow{\scriptstyle\iota} \\ G\times G & \xrightarrow[\mu_G]{} & G \end{array}$$

et que les applications $\iota \times \iota$ et μ_G sont différentiables, il en est de même de $\iota \circ \mu_H$, donc de μ_H.

Ceci prouve que *la variété invariante maximale H du sous-fibré $E\mathfrak{h}$ est un groupe de Lie, et, par suite, un sous-groupe du groupe de Lie G.*

On désignera ce sous-groupe par $G(\mathfrak{h})$.

Remarque 3. D'après le lemme 5 les raisonnements produits montrent aussi que *la sous-variété connexe H du groupe de Lie G sera son sous-groupe si l'ensemble H_{abstr} est un sous-groupe du groupe G_{abstr}.*

Comme l'espace tangent en e au sous-groupe $H = G(\mathfrak{h})$ est le sous-espace $E_e^{\mathfrak{h}} = \mathfrak{h}$, *l'algèbre de Lie du sous-groupe H est la sous-algèbre donnée $\mathfrak{h}$.*

Nous pouvons passer maintenant à la démonstration du théorème 1.

D é m o n s t r a t i o n d u t h é o r è m e 1. Nous savons déjà qu'à tout sous-groupe H du groupe de Lie G est associée la sous-algèbre $\mathfrak{h} = \mathfrak{l}(H)$ de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$, et à toute sous-algèbre $\mathfrak{h} \subset \mathfrak{g}$, un sous-groupe connexe $H = G(\mathfrak{h})$, et de plus $\mathfrak{l}(H) = \mathfrak{h}$. Pour achever la démonstration du théorème 1, il nous reste à prouver seulement que pour tout sous-groupe connexe H du groupe de Lie G, on a l'égalité $H = G(\mathfrak{h})$, où $\mathfrak{h} = \mathfrak{l}(H)$. Or, les sous-groupes H et $G(\mathfrak{h})$ ont la même algèbre de Lie $\mathfrak{h}$, donc ils sont confondus en tant que groupuscules (cf. leçon 7). Ceci signifie que les sous-groupes H et $G(\mathfrak{h})$ possèdent les mêmes voisinages de l'unité. Donc, $H = G(\mathfrak{h})$, puisque étant connexes, les groupes H et $G(\mathfrak{h})$ sont engendrés par chaque voisinage de l'unité. □

LEÇON 12

Autres définitions de la notion de sous-groupe d'un groupe de Lie.— Sous-groupes topologiques de groupes de Lie.— Sous-groupes fermés de groupes de Lie.— Groupes algébriques. — Groupes des automorphismes d'algèbres.— Groupes des automorphismes des groupes de Lie.— Idéaux et sous-groupes invariants.— Variétés quotients des groupes de Lie.— Groupes quotients de groupes de Lie.— Calcul des groupes fondamentaux.— Simple connexité des groupes SU(*n*) et Sp(*n*). — Groupe fondamental du groupe U(*n*).

La remarque 3 de la leçon précédente nous donne une définition de la notion de sous-groupe d'un groupe de Lie, formellement plus large. Il s'avère que les conditions imposées aux sous-groupes des groupes de Lie peuvent être relâchées dans d'autres directions.

On dit qu'une application différentiable $\Phi: N \to M$ est une *immersion* si pour tout point $a \in N$, l'application linéaire

$$(d\Phi)_a: \mathsf{T}_a(N) \to \mathsf{T}_{\Phi a}(M)$$

est un monomorphisme (est injective).

Proposition 1. *Tout monomorphisme* $\Phi: H \to G$ *de groupes de Lie est une immersion.*

D é m o n s t r a t i o n. Considérons l'application exponentielle

$$\exp: \mathfrak{l}(G) \to G.$$

Si les éléments de l'algèbre $\mathfrak{l}(G)$ sont traités comme des sous-groupes à un paramètre, cette application est définie par la formule

$$\exp \beta = \beta(1).$$

Comme $\mathfrak{l}(\Phi)\beta = \Phi \circ \beta$ (proposition 5 de la leçon 2), on en déduit le diagramme commutatif

$$\begin{array}{ccc} \mathfrak{l}(H) & \xrightarrow{\mathfrak{l}(\Phi)} & \mathfrak{l}(G) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ H & \xrightarrow[\Phi]{} & G \end{array}$$

Si l'application linéaire $\mathfrak{l}(\Phi) = (d\Phi)_e$ n'est pas injective, il existe dans un voisinage normal du zéro de l'algèbre $\mathfrak{l}(H) = \mathsf{T}_e(H)$ un vecteur A non nul appartenant au noyau de l'application $\mathfrak{l}(\Phi)$, c'est-à-dire tel que $\mathfrak{l}(\Phi)A = 0$. Mais alors

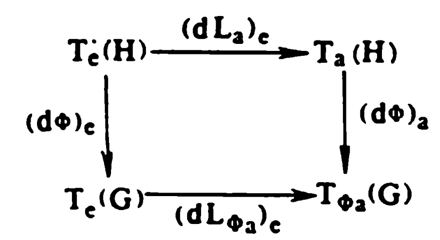
$$(\Phi \circ \exp)A = (\exp \circ \mathfrak{l}(\Phi))A = \exp 0 = e,$$

et, par suite, $\exp A = e$ (puisque l'application Φ est injective). Comme ceci est impossible (l'application exp est un difféomorphisme sur un voisinage normal), on en déduit que l'application $(d\Phi)_e$ est un monomorphisme.

Par ailleurs, le fait que l'application Φ est un homomorphisme signifie que pour tout élément $a \in H$, on a la relation

$$\Phi \circ L_a = L_{\Phi a} \circ \Phi.$$

Pour les différentielles cela signifie qu'on a le diagramme commutatif



d'où il s'ensuit (puisque les flèches horizontales sont des isomorphismes) que l'application $(d\Phi)_a$ est aussi un monomorphisme. □

Corollaire. *Un groupe de Lie H pour lequel H_{abstr} est un sous-ensemble de G_{abstr} et l'injection $\iota: H \to G$ un homomorphisme de groupes de Lie, est un sous-groupe du groupe de Lie G.*

Démonstration. En vertu de la proposition 1, l'application ι est une immersion, ce qui exprime que H est un sous-gronpe du groupe de Lie G. □

D'après ce corollaire, *les groupes de Lie de matrices introduits dans la leçon* 3 *ne sont autres que des sous-groupes du groupe de Lie* $\mathrm{GL}(n; \mathbb{R})$.

Les sous-groupes H du groupe de Lie G les plus intéressants sont manifestement ceux dont la topologie est induite par celle du groupe G, c'est-à-dire ceux pour lesquels le groupe topologique H_{top} est un sous-groupe du groupe topologique G_{top}. Nous les appellerons *sous-groupes topologiques* de groupes de Lie.

On rappelle qu'un sous-ensemble A d'un espace topologique X est par définition *localement fermé* si tout point $a \in A$ admet dans X un voisinage U tel que l'intersection $A \cap U$ est fermée dans U. Nous avons déjà eu affaire à cette notion dans la leçon 7 en étudiant les sous-groupuscules.

Nous aurons besoin du lemme suivant de la théorie des groupes topologiques:

Lemme 1. *Tout sous-groupe localement fermé H d'un groupe topologique G est fermé.*

D é m o n s t r a t i o n. Soit U un voisinage du point e tel que $H \cap U$ soit fermé dans U. Il est évident que sans nuire à la généralité on peut admettre que $U^{-1} = U$. Considérons un point quelconque $x \in \overline{H}$. Alors $xU \cap H \neq \varnothing$. Soit $y \in xU \cap H$. La translation à gauche $L_y : a \mapsto ya$ étant un homéomorphisme, l'ensemble $y(U \cap H)$ est fermé dans yU. Or, comme $y \in H$, il vient $y(U \cap H) = yU \cap H$, donc $yU \cap H$ est fermé dans yU, c'est-à-dire que $\overline{yU \cap H} \cap yU = yU \cap H$. Par ailleurs, $x \in yU^{-1} = yU$, donc $x \in yU \cap \overline{H} \subset \overline{yU \cap H}$. Par conséquent, $x \in yU \cap H$, et $x \in H$. $\square$

On dira qu'un sous-groupe H d'un groupe de Lie G est *fermé* si l'ensemble de ses points est un sous-ensemble fermé de G.

On remarquera que la topologie du sous-groupe H ne figure sous aucune forme dans cette définition. Il n'empêche toutefois qu'elle est fixée de façon unique par la condition de fermeture:

Proposition 2. *Un sous-groupe H d'un groupe de Lie G est fermé si et seulement si sa topologie est induite par celle du groupe G, c'est-à-dire si et seulement s'il est un sous-groupe topologique.*

D é m o n s t r a t i o n. D'après le lemme 1, nous prouverons la condition suffisante lorsque nous aurons établi que *toute sous-variété N dont la topologie est induite par celle de la variété ambiante M est localement fermée.* Or, ceci coule presque de source. En effet, tout point d'une telle sous-variété possède dans M un voisinage U qui coupe N suivant une sous-variété plate V dans U. D'autre part, il est clair que toute sous-variété plate est localement fermée.

Pour établir la condition nécessaire, nous utiliserons le fait que tout sous-groupe fermé H d'un groupe de Lie G est une sous-variété localement redressable admettant une base dénombrable.

Donc, il existe dans G un voisinage U de l'unité e tel que l'intersection $U \cap H$ est la réunion d'un nombre dénombrable de sous-variétés plates V_ξ, $\xi \in \Xi \subset \mathbb{R}^{n-m}$. Ceci étant, le sous-groupe H étant fermé, l'ensemble Ξ (qui est l'image de l'ensemble $U \cap H$ par l'application continue définie par la formule (2) de la leçon précédente) est localement fermé. Etant dénombrable, il possède donc au moins un point isolé ξ_0. La sous-variété plate correspondante V_{ξ_0} est telle que chacun de ses points a possède dans U un voisinage U_a tel que l'intersection $U_a \cap H$ est contenue dans V_{ξ_0}, et, par suite, est un voisinage du point a dans H. Cela signifie que le point a possède dans U (donc dans G) un système fondamental de voisinages découpant un système fondamental de voisinages du point a dans H. En appliquant la translation à gauche $L_{ba^{-1}}$, on trouve que cette propriété est caractéristique pour tout point $b \in H$. Mais alors la topologie de H sera, par définition, induite par celle de l'espace ambiant G. □

D'après le corollaire du lemme 1 de la leçon précédente (ou, si l'on veut, d'après le corollaire de la proposition 3 de la leçon 4), on déduit de la proposition 1 que *tout sous-groupe fermé H d'un groupe de Lie G est muni d'une seule structure de variété différentiable.*

Ce qui est étonnant c'est que seule la condition de fermeture suffit pour que H soit muni d'une structure différentiable qui en fait un groupe de Lie et un sous-groupe du groupe de Lie G.

Théorème 1 (Cartan). *Si un sous-ensemble fermé H d'un groupe de Lie G est un sous-groupe du groupe G_{abstr}, alors H est muni d'une seule structure différentiable (compatible avec la topologie induite sur H) pour laquelle H est un sous-groupe du groupe de Lie G (et, en particulier, est un groupe de Lie).*

D é m o n s t r a t i o n. Tout groupe de Lie G est en même temps groupuscule de Lie. Comme H est fermé, pour tout voisinage U du point e dans le groupuscule G, l'intersection $U \cap H$ est fermée dans U. Donc, H est un sous-groupuscule du groupuscule de Lie G (cf. définition 1 de la leçon 7). Par conséquent, d'après le théorème de Cartan pour les groupuscules (proposition 1 de la leçon 7), le groupuscule H est localement plat, c'est-à-dire, en d'autres termes, son intersection $U \cap H$ avec une carte U en e est une sous-variété plate de cette carte de la forme $\exp \mathfrak{h}$, où $\mathfrak{h}$ est une sous-algèbre de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$. Ceci signifie que H est localement confondu en e avec le sous-groupe connexe $G(\mathfrak{h})$ du groupe de Lie G. Comme tout voisinage d'un groupe topologique connexe engendre le groupe tout entier (lemme 3 de la leçon 9), ceci prouve que le groupe $G(\mathfrak{h})$ (plus exactement, le groupe $G(\mathfrak{h})_{\mathrm{abstr}}$ muni de la topologie induite) est confondu avec la composante de l'unité H_e

du groupe H. Ceci munit le groupe H_e, donc le groupe H tout entier, d'une structure différentiable pour laquelle ce groupe est un sous-groupe du groupe de Lie G.

D'après la remarque faite ci-dessus, cette structure différentiable est unique, puisque le sous-groupe H est fermé. □

A noter qu'étant fermé, le sous-groupe H de G sera sous-groupe topologique de G.

Le théorème 1 est un puissant outil de reconnaissance des groupes de Lie.

Exemple. On dit qu'un sous-groupe du groupe $\mathrm{GL}(n;\mathbb{R})$ (resp. $\mathrm{GL}(n;\mathbb{C})$) est un *groupe algébrique* s'il est l'intersection du groupe $\mathrm{GL}(n;\mathbb{R})$ (resp. $\mathrm{GL}(n;\mathbb{C})$) avec une variété algébrique de l'espace $\mathbb{R}(n)=\mathbb{R}^{n^2}$ (resp. de l'espace $\mathbb{C}(n)=\mathbb{C}^{n^2}$), c'est-à-dire avec l'ensemble des zéros communs d'un système de polynômes en n^2 indéterminées.

Comme toute variété algébrique est visiblement fermée, il vient, en vertu du théorème 1, que *tout groupe algébrique est un groupe de Lie de matrices.*

Ceci prouve immédiatement que tous les groupes de la leçon 1 ($\mathrm{SL}(n)$, $\mathrm{O}(n)$, $\mathrm{Sp}(n)$, $\mathrm{U}(n)$, etc.) sont des groupes de Lie (à noter que le groupe $\mathrm{U}(n)$ doit être traité non pas comme un sous-groupe du groupe $\mathrm{GL}(n;\mathbb{C})$, mais, en vertu de l'inclusion $\mathrm{GL}(n;\mathbb{C})\subset\subset\mathrm{GL}(2n;\mathbb{R})$, comme un sous-groupe du groupe $\mathrm{GL}(2n;\mathbb{R})$).

Soit $\mathcal{A}$ une algèbre de dimension finie (généralement non associative et non de Lie) sur le corps $\mathbb{R}$ ou $\mathbb{C}$, et soit $e_1, \ldots, e_n$ une de ses bases. Il est évident qu'une application linéaire inversible $\Phi\colon \mathcal{A}\to\mathcal{A}$ est un automorphisme de l'algèbre $\mathcal{A}$ si et seulement si $\Phi(e_ie_j)=\Phi(e_i)\,\Phi(e_j)$ pour tous $i, j=1, \ldots, n$. Par conséquent, si $\Phi(e_i)=x_i^je_j$ et $e_ie_j=c_{ij}^ke_k$ et donc

$$\Phi\,(e_ie_j)=\Phi(c_{ij}^ke_k)=c_{ij}^kx_k^le_l,$$

$$\Phi(e_i)\,\Phi(e_j)=(x_i^pe_p)\,(x_j^qe_q)=c_{pq}^lx_i^px_j^qe_l,$$

alors Φ est un automorphisme si et seulement si

$$c_{ij}^kx_k^l=c_{pq}^lx_i^px_j^q$$

quels que soient $i, j, l=1, \ldots, n$. Ceci signifie que les matrices (x_i^j) associées aux automorphismes de l'algèbre $\mathcal{A}$ forment un groupe algébrique, donc différentiable. Donc, la donnée d'une base $e_1, \ldots$ $\ldots, e_n$ définit un isomorphisme du groupe des automorphismes $\mathrm{Aut}\,\mathcal{A}$ de l'algèbre $\mathcal{A}$ sur un groupe de Lie algébrique de matrices. La structure de groupe de Lie transportée sur $\mathrm{Aut}\,\mathcal{A}$ au moyen de

cet isomorphisme ne dépend pas de toute évidence du choix de la base $e_1, \ldots, e_n$. Donc, *le groupe* Aut $\mathcal{A}$ *des automorphismes d'une algèbre* $\mathcal{A}$ *de dimension finie est un groupe de Lie.*

Trouvons l'algèbre de Lie de ce groupe.

Proposition 2. *L'algèbre de Lie du groupe* Aut $\mathcal{A}$ *est l'algèbre de Lie* Der $\mathcal{A}$ *de toutes les dérivations de l'algèbre* $\mathcal{A}$:

$$\mathfrak{l}(\mathrm{Aut}\,\mathcal{A}) = \mathrm{Der}\,\mathcal{A}.$$

D é m o n s t r a t i o n. En choisissant une base dans $\mathcal{A}$, on peut considérer que le groupe Aut $\mathcal{A}$ et l'algèbre de Lie Der $\mathcal{A}$ sont composés de matrices.

Soit $D \in \mathrm{Der}\,\mathcal{A}$. Alors, pour tous éléments x, $y \in \mathcal{A}$ et tout $p \geqslant 0$, on aura l'égalité

$$D^p(xy) = \sum_{i=0}^{p} \binom{p}{i} D^i x \cdot D^{p-i} y \quad (\textit{formule de Leibnitz})$$

et, par suite,

$$(e^{tD})(xy) = \sum_{p=0}^{\infty} \frac{t^p}{p!} D^p(xy) =$$

$$= \sum_{p=0}^{\infty} \sum_{i=0}^{p} \frac{1}{p!} \binom{p}{i} t^p \cdot D^i x \cdot D^{p-i} y =$$

$$= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{t^{i+j}}{i!\,j!} D^i x \cdot D^j y =$$

$$= \left(\sum_{i=0}^{\infty} \frac{t^i}{i!} D^i x \right) \left(\sum_{j=0}^{\infty} \frac{t^j}{j!} D^j x \right) = e^{tD} x \cdot e^{tD} y$$

(la convergence de toutes ces séries s'acquiert par des calculs classiques à l'aide des normes matricielles). Ceci signifie que $e^{tD} \in \mathrm{Aut}\,\mathcal{A}$, donc que $D \in \mathfrak{l}(\mathrm{Aut}\,\mathcal{A})$.

Réciproquement, soit $D \in \mathfrak{l}(\mathrm{Aut}\,\mathcal{A})$, c'est-à-dire que $e^{tD} \in$ $\in \mathrm{Aut}\,\mathcal{A}$. Alors

$$(e^{tD} - E)(xy) = e^{tD} x \cdot e^{tD} y - x \cdot y =$$

$$= (e^{tD} x - x)\, e^{tD} y + x\,(e^{tD} y - y),$$

donc,

$$D(xy) = \lim_{t\to 0} \frac{(e^{tD} - E)(xy)}{t} =$$
$$= \lim_{t\to 0} \frac{e^{tD}x - x}{t} e^{tD}y + \lim_{t\to 0} x \frac{e^{tD}y - y}{t} =$$
$$= Dx \cdot y + x \cdot Dy.$$

Et, par suite, $D \in \operatorname{Der} \mathcal{A}$. □

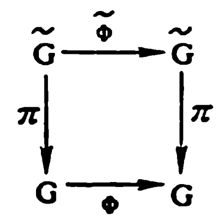
Supposons, en particulier, que $\mathcal{A}$ est l'algèbre de Lie $\mathfrak{g}$ d'un groupe de Lie connexe G. Comme le foncteur de Lie est fidèle (cf. leçon 10) sur la catégorie des groupes de Lie simplement connexes, tout automorphisme $\mathfrak{g} \to \mathfrak{g}$ de l'algèbre de Lie $\mathfrak{g}$ est réalisé par un automorphisme $G \to G$ du groupe de Lie G. Ceci montre que *le groupe des automorphismes* Aut G *d'un groupe de Lie simplement connexe* G *est isomorphe au groupe des automorphismes* Aut $\mathfrak{g}$ *de son algèbre de Lie* $\mathfrak{g}$:

$$\operatorname{Aut} G \approx \operatorname{Aut} \mathfrak{g}.$$

En transportant la structure différentiable de Aut $\mathfrak{g}$ à Aut G au moyen de cet isomorphisme, on munit Aut G d'une structure de groupe de Lie. Ceci étant, d'après ce qui a été dit plus haut, *l'algèbre de Lie du groupe de Lie* Aut G *sera l'algèbre de Lie* Der $\mathfrak{g}$:

$$\mathfrak{l}(\operatorname{Aut} G) = \operatorname{Der} \mathfrak{l}(G).$$

Pour obtenir le même résultat pour un groupe de Lie connexe G, considérons son revêtement universel $\tilde{G}$. On sait (cf. leçon 9) que $\tilde{G}$ dépend fonctoriellement de G et, par suite, tout automorphisme $\Phi: G \to G$ définit un seul automorphisme $\tilde{\Phi}: \tilde{G} \to \tilde{G}$ tel que l'on a le diagramme commutatif



Ceci définit l'application (qui est visiblement un monomorphisme)

$$\operatorname{Aut} G \to \operatorname{Aut} \tilde{G}.$$

L'image de ce monomorphisme est composée d'automorphismes $\Psi: \tilde{G} \to \tilde{G}$ pour lesquels l'égalité $\pi(a) = \pi(b)$ entraîne l'égalité $\pi\Psi(a) = \pi\Psi(b)$. Cette condition revient exactement à dire que l'automorphisme Ψ envoie dans lui-même le noyau $K = \operatorname{Ker} \pi$

du revêtement π. Donc, l'ensemble de ces automorphismes est fermé, et, par suite, (théorème 1) est un sous-groupe du groupe de Lie Aut $\tilde{G}$, donc un groupe de Lie. En transportant cette structure de groupe de Lie sur le groupe Aut G, on fait de ce dernier un groupe de Lie.

Nous avons ainsi prouvé que *le groupe des automorphismes d'un groupe de Lie connexe est un groupe de Lie.*

Revenons maintenant aux groupes de Lie et à leurs sous-groupes fermés. Comme l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$ définit de façon unique le groupe de Lie simplement connexe G, toutes les sous-algèbres $\mathfrak{h}$ de l'algèbre de Lie $\mathfrak{g}$ se répartissent en deux classes: aux sous-algèbres d'une classe sont associés les sous-groupes fermés du groupe de Lie simplement connexe G, et aux sous-algèbres de l'autre classe, les sous-groupes non fermés. Ceci pose le problème de la caractérisation algébrique intérieure des sous-algèbres de ces classes, ou, ce qui revient au même, des sous-groupes correspondants. Nous n'étudierons pas ce problème dans le cas général, nous contentant du résultat suivant, semble-t-il le plus intéressant et le plus inattendu:

Théorème 2. *Tout sous-groupe invariant connexe H d'un groupe de Lie simplement connexe G est fermé.*

Prouvons tout d'abord une proposition caractérisant les sous-algèbres de Lie associées aux sous-groupes invariants:

Proposition 3. *L'algèbre de Lie $\mathfrak{l}(H)$ de tout sous-groupe invariant H d'un groupe de Lie G est un idéal de l'algèbre de Lie $\mathfrak{l}(G)$ du groupe G. Réciproquement, si un sous-groupe H est connexe et l'algèbre de Lie $\mathfrak{l}(H)$ est un idéal de l'algèbre de Lie $\mathfrak{l}(G)$, alors le sous-groupe H est invariant.*

Démonstration. Cette proposition est complètement analogue à la proposition 2 de la leçon 7 et peut être prouvée exactement de la même manière. Mais pour varier nous allons produire une autre démonstration basée sur la proposition 2 de la leçon 7.

Il est évident que, d'après cette proposition, il suffit de montrer qu'*un sous-groupe connexe H d'un groupe de Lie G est invariant si et seulement si un voisinage quelconque V de son unité est un sous-groupuscule invariant du groupe G* (considéré comme un groupuscule). En effet, si H est invariant, V l'est par définition. Réciproquement, supposons que V est invariant et soit $g \in G$. Il existe alors dans V un voisinage W de l'unité tel que $g^{-1}Wg \subset V$. Comme H est un groupe connexe, ses éléments a sont de la forme $a_1a_2 \dots a_k$, où $a_1, a_2, \dots, a_k \in W$. Donc,

$$g^{-1}ag = g^{-1}a_1g \cdot g^{-1}a_2g \cdot \dots \cdot g^{-1}a_kg \in V \cdot V \cdot \dots \cdot V \subset H.$$

Donc, le sous-groupe H est invariant. □

A noter qu'il existe des sous-groupes H non invariants et non connexes pour lesquels la sous-algèbre $\mathfrak{l}(H)$ est un idéal de l'algèbre de Lie $\mathfrak{l}(G)$. Leurs composantes de l'unité sont invariantes.

La proposition suivante est un complément essentiel à la proposition 3:

Proposition 4. *Le noyau* Ker Φ *de tout homomorphisme* $\Phi: G \to H$ *de groupes de Lie est un sous-groupe du groupe de Lie* G. *L'algèbre de Lie* $\mathfrak{l}(\mathrm{Ker}\ \Phi)$ *est confondue avec le noyau* Ker $\mathfrak{l}(\Phi)$ *de l'application induite* $\mathfrak{l}(\Phi): \mathfrak{l}(G) \to \mathfrak{l}(H)$ *d'algèbres de Lie*:

$$\mathfrak{l}(\mathrm{Ker}\ \Phi) = \mathrm{Ker}\ \mathfrak{l}(\Phi).$$

D é m o n s t r a t i o n. Le noyau Ker Φ étant fermé, la première affirmation résulte du théorème 1. Tout groupe à un paramètre $\beta: \mathbb{R} \to \mathrm{Ker}\ \Phi$ de ce noyau est envoyé par l'homomorphisme Φ dans une application constante, c'est-à-dire dans le zéro de l'algèbre $\mathfrak{l}(H)$. Donc, $\mathfrak{l}(\mathrm{Ker}\ \Phi) \subset \mathrm{Ker}\ \mathfrak{l}(\Phi)$. Réciproquement, si un sous-groupe à un paramètre $\beta: \mathbb{R} \to G$ du groupe G est contenu dans $\mathrm{Ker}\ \mathfrak{l}(\Phi)$, alors $\Phi \circ \beta = \mathrm{const}$, c'est-à-dire que $\beta(t) \in \mathrm{Ker}\ \Phi$ pour tous les $t \in \mathbb{R}$ et, par suite, $\beta \in \mathfrak{l}(\mathrm{Ker}\ \Phi)$. Donc, $\mathrm{Ker}\ \mathfrak{l}(\Phi) \subset$ $\subset \mathfrak{l}(\mathrm{Ker}\ \Phi)$.

Nous pouvons maintenant démontrer le théorème 2.

D é m o n s t r a t i o n d u t h é o r è m e 2. Soient $\mathfrak{g} =$ $= \mathfrak{l}(G)$ et $\mathfrak{h} = \mathfrak{l}(H)$. Le sous-groupe H étant invariant, la sous-algèbre $\mathfrak{h}$ est un idéal, et donc est définie l'algèbre quotient $\mathfrak{g}/\mathfrak{h}$. D'après le théorème 1 de la leçon 10 (à noter que pour l'instant ce théorème n'a été prouvé que sous l'hypothèse de la véracité du théorème d'Ado), il existe un groupe de Lie simplement connexe N dont l'algèbre de Lie est $\mathfrak{g}/\mathfrak{h}$. Comme le foncteur de Lie est fidèle sur la catégorie des groupes de Lie simplement connexes, il existe un homomorphisme $\Phi: G \to N$ de groupes de Lie réalisant l'homomorphisme canonique $\varphi: \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$, c'est-à-dire tel que $\mathfrak{l}(\Phi) = \varphi$. En vertu de la proposition 4, le noyau Ker Φ est un sous-groupe fermé dont l'algèbre de Lie est

$$\mathfrak{l}(\mathrm{Ker}\ \Phi) = \mathrm{Ker}\ \mathfrak{l}(\Phi) = \mathrm{Ker}\ \varphi = \mathfrak{h}.$$

La composante de l'unité $(\mathrm{Ker}\ \Phi)_e$ de ce noyau est fermée aussi et son algèbre de Lie est aussi l'idéal $\mathfrak{h}$. Donc, le groupe de Lie G contient deux sous-groupes connexes H et $(\mathrm{Ker}\ \Phi)_e$ ayant la même algèbre de Lie. Donc, $H = (\mathrm{Ker}\ \Phi)_e$, en vertu du théorème 1 de la leçon précédente. Et, par suite, le sous-groupe H est fermé. □

Remarque 1. La proposition 3 dit que la correspondance de Lie associe biunivoquement les sous-groupes invariants connexes du groupe de Lie G aux idéaux de l'algèbre de Lie $\mathfrak{g}$. Il est évident que

dans ces conditions *si le groupe G est le produit direct $A \times B$ de deux sous-groupes invariants A et B, l'algèbre de Lie $\mathfrak{g}$ sera la somme directe $\mathfrak{a} \oplus \mathfrak{b}$ des idéaux $\mathfrak{a} = \mathfrak{l}(A)$ et $\mathfrak{b} = \mathfrak{l}(B)$.* Mais la réciproque est généralement fausse même pour des groupes de Lie connexes G. Si $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$, et A et B sont des sous-groupes invariants d'un groupe connexe G tels que $\mathfrak{l}(A) = \mathfrak{a}$ et $\mathfrak{l}(B) = \mathfrak{b}$, le groupe G n'est pas nécessairement le produit direct $A \times B$ des groupes A et B. On peut tout au plus affirmer que le groupe G est engendré par les groupes A et B (puisque tout élément d'un voisinage normal arbitraire est visiblement le produit d'éléments des groupes A et B, et le groupe G, étant connexe, est engendré par ce voisinage) et que l'intersection $A \cap B$ sera un sous-groupe invariant de dimension zéro, puisque c'est un sous-groupe du groupe de Lie G ayant une algèbre de Lie nulle. En particulier, si l'intersection $A \cap B$ est fermée (ce qui, en vertu du théorème 2, a toujours lieu si le groupe G est simplement connexe), elle est discrète. De plus, si les sous-groupes A et B sont connexes, alors étant continue, l'application $(a, b) \mapsto aba^{-1}b^{-1}$ de la variété connexe $A \times B$ dans la variété discrète $A \cap B$ envoie la variété $A \times B$ tout entière dans le point $e \in G$, c'est-à-dire que les sous-groupes A et B commutent. Donc, l'application $A \times B \to G$ définie par la formule $(a, b) \mapsto ab$ est un épimorphisme. Comme cet épimorphisme induit de toute évidence l'isomorphisme identique $\mathfrak{a} \oplus \mathfrak{b} \to \mathfrak{g}$ d'algèbres de Lie, la proposition 4 nous dit que son noyau est discret et, par suite, c'est un revêtement de groupe, donc un isomorphisme pour le groupe simplement connexe G. Ceci prouve qu'*un groupe de Lie simplement connexe G se décompose en le produit direct $A \times B$ de sous-groupes connexes A et B si et seulement si son algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$ se décompose en la somme directe $\mathfrak{a} \oplus \mathfrak{b}$ des sous-algèbres $\mathfrak{a} = \mathfrak{l}(A)$ et $\mathfrak{b} = \mathfrak{l}(B)$.*

Pour tout sous-groupe H d'un groupe de Lie G, les composantes des classes à gauche aH, $a \in G$, ne sont évidemment autres que les variétés invariantes maximales du sous-fibré $E^{\mathfrak{h}}$, $\mathfrak{h} = \mathfrak{l}(H)$. Il existe donc dans le groupe G une carte (cubique) $(U, h) = (U, x^1, \ldots, x^n)$ telle que $e \in U$ et pour tout $a \in G$, l'intersection $U \cap aH$ est la réunion de sous-variétés plates de la forme V_ξ, $\xi \in \mathbb{R}^{n-m}$. Cette propriété caractérise, en particulier, l'intersection $U \cap H$. Si le sous-groupe H est fermé (ce que nous supposerons toujours dans la suite), la carte (U, h) peut être choisie de telle sorte que l'intersection $U \cap H$ soit composée d'une seule sous-variété V_0 (voir ci-dessus la démonstration de la proposition 1). Soit W un voisinage (cubique relativement à h) du point e de G tel que $W^{-1} \subset W$ et $W^4 \subset U$. Si des points $a, b \in W$ sont congrus modulo H, c'est-à-dire si $a^{-1}b \in H$, alors $a^{-1}b \in W^2 \cap H \subset U \cap H = V_0$, i.e. $b \in aV_0$. Comme $a \in aV_0$ et aV_0 est connexe avec V_0, ceci prouve

que les points a et b appartiennent à une même composante à gauche suivant H de l'intersection $U \cap aH$, c'est-à-dire à une même variété V_ξ. Comme la réciproque est évidente (si $a, b \in W \cap V_\xi$, alors $a^{-1}b \in H$), ceci prouve que les points $a, b \in W$ appartiennent à la même classe à gauche suivant H si et seulement s'il existe un ξ tel que $a, b \in V_\xi$, ou, en d'autres termes, *les intersections* $V_\xi \cap W$ *pour des* ξ *différents appartiennent à des classes à gauche suivant* H *différentes.*

Considérons maintenant l'ensemble G/H de toutes les classes à gauche aH, $a \in G$. Le groupe G opère sur cet ensemble selon la formule

$$g(aH) = (ga)H, \quad g \in G, \quad aH \in G/H.$$

L'application $aH \to g(aH)$ sera désignée par $\overline{L}_g$. Elle est reliée à la translation à gauche L_g dans le groupe G par la formule $\overline{L}_g \circ \omega = \omega \circ L_g$, où ω est l'application canonique $G \to G/H$, $a \to aH$.

Soit $\overline{W}$ l'image d'un voisinage W par l'application ω.

D'après ce qui a été dit plus haut, l'application $\overline{h}\colon \overline{W} \to \mathbb{R}^{n-m}$ qui à toute classe $aH \in \overline{W}$ associe un point $\xi \in \mathbb{R}^{n-m}$ pour lequel $a \in V_\xi \cap W$ est défini de façon unique. Comme elle est visiblement injective et que l'ensemble $h(\overline{W})$ est un cube ouvert de $\mathbb{R}^{n-m}$ de demi-largeur c égale à celle du voisinage W, le couple $(\overline{W}, \overline{h})$ est une carte de G/H contenant le point H. Cette carte est reliée à la carte (W, h) par le diagramme commutatif

$$(1) \qquad \begin{array}{ccc} W & \longrightarrow & \overline{W} \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle \overline{h}} \\ \mathbb{R}^n & \longrightarrow & \mathbb{R}^{n-m} \end{array}$$

dont la flèche horizontale supérieure est l'application $\omega\colon a \to aH$ et l'inférieure, la projection parallèlement aux m premiers axes de coordonnées.

En associant à chaque classe $aH \in \overline{W}$, $a \in W$, le point de W de coordonnées

$$x^1 = 0, \ldots, x^m = 0, \quad x^{m+1} = x^{m+1}(a), \ldots, x^n = x^n(a),$$

on obtient une application $\sigma\colon \overline{W} \to W$, section de l'application $\omega\colon \overline{W} \to W$, c'est-à-dire telle que $\omega \circ \sigma = \text{id}$ sur $\overline{W}$. La section σ est reliée aux applications h et $\overline{h}$ par la formule

$$\overline{h} = h \circ \sigma.$$

On remarquera que l'application $\sigma \circ \omega\colon \overline{W} \to W$ est différentiable.

L'application $\overline{L}_a$ étant bijective pour tout $a \in G$, le couple $(a\overline{W}, \overline{h}_a)$, où $\overline{h}_a = \overline{h} \circ \overline{L}_a$ est une carte en $aH \in G/H$. Si $a\overline{W} \cap b\overline{W} \neq \varnothing$, alors sur $h_a\,(a\overline{W} \cap b\overline{W})$ l'application $\overline{h}_b \circ \overline{h}_a^{-1}$ est justiciable de la formule

$$\overline{h}_b \circ \overline{h}_a^{-1} = \overline{h} \circ \overline{L}_b \circ \overline{L}_a^{-1} \circ \overline{h}^{-1} =$$

$$= h \circ \sigma \circ \overline{L}_{ba^{-1}} \circ \omega \circ h^{-1} = h \circ \sigma \circ \omega \circ L_{ba^{-1}} \circ h^{-1},$$

d'ou il s'ensuit aussitôt que cette application est différentiable et, par suite, la carte $(a\overline{W}, \overline{h}_a)$ est compatible avec la carte $(b\overline{W}, \overline{h}_b)$.

Comme les cartes $(a\overline{W}, \overline{h}_a)$ recouvrent de toute évidence G/H, ceci prouve qu'elles forment un atlas et donc définissent une structure différentiable sur G/H. Muni de cette structure l'ensemble G/H s'appelle *variété quotient* (ou *espace homogène*) du groupe de Lie G par son sous-groupe fermé H.

L'application $\omega\colon G \to G/H$ envoie les cartes aW dans les cartes $a\overline{W}$ et, par suite, (cf. diagramme (1)) est une projection $\mathbb{R}^n \to \mathbb{R}^{n-m}$ parallèlement aux axes de coordonnées correspondants. Ceci signifie que *l'application canonique*

$$\omega : G \to G/H, \quad a \mapsto aH,$$

est différentiable et possède le même rang $n - m$ *en tous les points* (de sorte que sa différentielle $(d\omega)_a$ en tout point $a \in G$ est un épimorphisme).

La section $\sigma\colon \overline{W} \to W$ est de toute évidence une application différentiable.

L'application

$$\overline{\mu}\colon G \times G/H \to G/H, \quad (g, aH) \mapsto (ga)H,$$

qui définit l'action du groupe G sur G/H est reliée à la multiplication dans le groupe G par le diagramme commutatif

$$\begin{array}{ccc} G\times G & \xrightarrow{\ \mu\ } & G \\ {\scriptstyle \mathrm{id}\times\omega}\downarrow & & \downarrow{\scriptstyle \omega} \\ G\times G/H & \xrightarrow[\ \pi\ ]{} & G/H \end{array}$$

Pour tous points $g, a \in G$, la restriction de $\mathrm{id} \times \omega$ au voisinage $gW \times a\overline{W}$ du point (g, aH) admet la section $\mathrm{id} \times \sigma$. Donc, la restriction de l'application $\overline{\mu}$ à ce voisinage est la composée des applications différentiables $\mathrm{id} \times \sigma$, μ, ω et, par suite, est différentiable. Ceci prouve que ***l'application*** $\overline{\mu}$ ***est différentiable.***

14—01642

Supposons maintenant que le sous-groupe H est invariant. La formule

$$aH \cdot bH = abH$$

définit une seule multiplication dans G/H, multiplication qui munit G/H d'une structure de groupe. Il est immédiat de voir que cette multiplication définit une application différentiable

$$G/H \times G/H \to G/H, \tag{2}$$

c'est-à-dire que *le groupe quotient G/H est un groupe de Lie.* En effet, soient a, $b \in G$. Considérons dans G/H les voisinages $a\overline{W}$ et $b\overline{W}$ des classes aH et bH. La restriction de l'application (2) au voisinage $a\overline{W} \times b\overline{W}$ est la composée des applications différentiables $\sigma \times \sigma$, μ et ω, donc elle est différentiable. Par conséquent, elle est partout différentiable. □

Proposition 5. *L'algèbre de Lie du groupe quotient G/H est isomorphe à l'algèbre quotient de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$ par l'idéal $\mathfrak{h} = \mathfrak{l}(H)$:*

$$\mathfrak{l}(G/H) \approx \mathfrak{g}/\mathfrak{h}.$$

D é m o n s t r a t i o n. Etant différentiable, l'application $\omega: G \to G/H$ est un homomorphisme de groupes de Lie et, par suite, induit l'homomorphisme

$$\mathfrak{l}(\omega): \mathfrak{g} \to \mathfrak{l}(G/H)$$

de leurs algèbres de Lie. Si les algèbres de Lie $\mathfrak{g}$ et $\mathfrak{l}(G/H)$ sont traitées comme des espaces tangents en e à G et G/H, l'homomorphisme $\mathfrak{l}(\omega)$ n'est autre que la différentielle

$$(d\omega)_e: \mathsf{T}_e(G) \to \mathsf{T}_e(G/H)$$

de l'application ω. Mais nous avons vu plus haut que cette différentielle était un épimorphisme de noyau $\mathfrak{h} = \mathsf{T}_e(H)$. Donc, l'homomorphisme $\mathfrak{l}(\omega)$ induit un isomorphisme de l'algèbre quotient $\mathfrak{g}/\mathfrak{h}$ sur l'algèbre $\mathfrak{l}(G/H)$. □

Signalons que le sous-groupe H est supposé fermé dans ce théorème. Les algèbres quotients $\mathfrak{g}/\mathfrak{h}$ par les idéaux $\mathfrak{h} \subset \mathfrak{g}$ auxquels sont associés les sous-groupes invariants non fermés du groupe G ne sont les algèbres de Lie d'aucun groupe quotient du groupe G (tout au moins si les groupes quotients sont pris dans leur acception habituelle).

Dans le cas où le sous-groupe invariant H n'est pas connexe, il résulte de la proposition 5 que les groupes quotients G/H et G/H_e, où H_e est la composante de l'unité du sous-groupe H (qui est visiblement invariante aussi) sont localement isomorphes. On peut pré-

ciser cette remarque en observant que puisque $H_e \subset H$, chaque classe suivant H_e est contenue dans une classe suivant H définie de façon unique, ce qui détermine l'application canoniqu

$$\rho: G/H_e \to G/H.$$

Les applications canoniques $\omega: G \to G/H$ et $\omega_e: G \to G/H_e$ étant continues et ouvertes, il en est de même de ρ. Si le sous-groupe H est invariant, l'application ρ est visiblement un homomorphisme.

Proposition 6. *Pour tout sous-groupe fermé H d'un groupe de Lie connexe G, l'application canonique*

$$\rho: G/H_e \to G/H$$

est un revêtement.

Démonstration. Le sous-groupe H_e étant ouvert dans H, l'unité du groupe G possède un voisinage connexe V tel que $V^{-1}V \subset \subset H_e$. Il est évident que pour prouver la proposition 6, il faut montrer que le voisinage $V(g) = \omega(gV)$ de tout point $\omega(g) = gH$ de la variété G/H est revêtu sans pli par l'application ρ.

Choisissons un représentant h_α dans chaque composante H_α du sous-groupe H et considérons les ensembles connexes ouverts $\omega_e(gVh_\alpha)$ de G/H_e. Ces ensembles sont disjoints (si $\omega_e(gvh_\alpha) = = \omega_e(gv'h_\beta)$, c'est-à-dire si $gvh_\alpha = gv'h_\beta h$, où $h \in H_e$, alors $h_\alpha = = v^{-1}v' \cdot h_\beta h$, où $v^{-1}v' \in H_e$, et, par suite, $h_\alpha = h_\beta$), leur somme est l'ensemble $\rho^{-1}V(g)$ (puisque les éléments de l'ensemble $V(g)$ sont les classes à gauche gvH, où $v \in V$, l'ensemble $\rho^{-1}V(g)$ est composé des classes à gauche $gvh_\alpha H_e$) et l'application ρ est une application bijective — donc un homéomorphisme — de chacun de ces ensembles sur $V(g)$ (si $\omega(gvh_\alpha) = \omega(gv'h_\alpha)$, alors $v'h_\alpha = vh_\alpha h$, où $h \in H$, et, par suite, $v^{-1}v' \in H$ et donc $v^{-1}v' \in H_e$; mais alors $h = h_\alpha^{-1} \cdot \cdot v^{-1}v' \cdot h_\alpha \in H_e$, car H_e est invariant dans H et, par suite, $\omega_e(gvh_\alpha) = = \omega_e(gv'h_\alpha)$). Par conséquent, le voisinage $V(g)$ est revêtu sans pli par l'application ρ. □

Corollaire. *Si dans les hypothèses de la proposition* 6, *la variété quotient G/H est simplement connexe, le sous-groupe H est connexe.* □

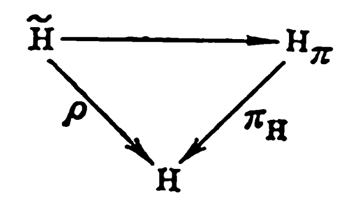
Nous pouvons prouver maintenant que la méthode générale d'établissement de la connexité de groupes topologiques, basée sur le lemme 2 de la leçon 1, peut être utilisée pour l'établissement de la simple connexité:

Proposition 7. *Si un groupe de Lie connexe G contient un sous-groupe connexe et simplement connexe fermé H tel que la variété quotient G/H soit simplement connexe, alors le groupe G est aussi simplement connexe.*

On a la proposition plus générale encore:

Proposition 8. *Pour tout sous-groupe connexe fermé H d'un groupe de Lie connexe G, tel que la variété quotient G/H soit simplement connexe, le groupe fondamental $\pi_1 G$ du groupe de Lie G est le groupe quotient du groupe fondamental $\pi_1 H$ du groupe de Lie H.*

Démonstration. Soit $\tilde{G}$ un groupe de Lie simplement connexe revêtant le groupe G, et soit $\pi: \tilde{G} \to G$ le revêtement correspondant. L'image réciproque $H_\pi = \pi^{-1}H$ du sous-groupe H par l'homomorphisme π est un sous-groupe fermé de $\tilde{G}$ et, de plus, l'application $\tilde{G}/H_\pi \to G/H$ induite par l'homomorphisme π est de toute évidence un difféomorphisme. Donc, la variété $\tilde{G}/H_\pi$ est aussi simplement connexe et, par suite, le sous-groupe H_π est connexe. Donc, la restriction π_H de l'application π à H_π est un revêtement $H_\pi \to H$ et, par suite, on a le diagramme commutatif



où $\rho: \check{H} \to H$ est un revêtement universel du groupe de Lie H. Ceci étant, l'application $H_\pi \to H$ (qui est aussi un revêtement) est un épimorphisme et, par suite, induit un épimorphisme du noyau $\operatorname{Ker} \rho$ de l'homomorphisme ρ sur le noyau $\operatorname{Ker} \pi_H$ de l'homomorphisme π_H. Ceci prouve la proposition 8 (et la proposition 7), puisque par définition $\pi_1 H = \operatorname{Ker} \rho$ et $\pi_1 G = \operatorname{Ker} \pi = \operatorname{Ker} \pi_H$. □

A titre d'exemple d'application de la proposition 7, considérons le groupe $\mathrm{SU}(n)$ des matrices unitaires unimodulaires. Comme (cf. leçon 1) pour tout $n > 1$ la variété quotient $\mathrm{SU}(n)/\mathrm{SU}(n-1) = \mathrm{U}(n)/\mathrm{U}(n-1)$ est difféomorphe à la sphère S^{2n-1}, alors en vertu du corollaire du lemme 1 de la leçon 9, cette variété quotient est simplement connexe. Le groupe $\mathrm{SU}(1)$ étant un singleton, donc simplement connexe, on en déduit en vertu de la proposition 7 par une récurrence évidente que *le groupe* $\mathrm{SU}(n)$ *est simplement connexe pour tout* $n \geqslant 1$.

Par un raisonnement analogue utilisant la simple connexité de la sphère $S^{4n-1} = \mathrm{Sp}(n)/\mathrm{Sp}(n-1)$ et du groupe $\mathrm{Sp}(1) \approx \mathbb{S}^3$, on établit que *le groupe* $\mathrm{Sp}(n) = \mathrm{U}^{\mathbf{H}}(n)$ *est simplement connexe pour tout* $n \geqslant 1$.

Dans le cas du groupe $\mathrm{U}(n)$, il faut appliquer la proposition 8, puisque le groupe $\mathrm{U}(1)$ qui est le groupe $\mathbb{S}^1$ des nombres complexes $|z| = 1$, n'est pas simplement connexe, car il admet un revêtement

non trivial $\mathbb{R} \to S^1$ défini par la formule $t \mapsto e^{2\pi i t}$, $t \in \mathbb{R}$. Ce revêtement étant universel, puisque $\mathbb{R}$ est simplement connexe, il s'ensuit que le groupe $\pi_1 S^1$ est isomorphe à son noyau, c'est-à-dire au groupe des entiers $\mathbb{Z}$. Donc, la même récurrence que ci-dessus mais utilisant la proposition 8 au lieu de la proposition 7, montre que le groupe fondamental $\pi_1 \mathrm{U}(n)$ du groupe de Lie $\mathrm{U}(n)$ est un groupe quotient du groupe de $\mathbb{Z}$.

Pour calculer entièrement ce groupe quotient, on se servira de la proposition suivante qui dans un certain sens est duale de la proposition 8:

Proposition 9. *Pour tout sous-groupe invariant connexe fermé H d'un groupe de Lie connexe G, le groupe fondamental $\pi_1 G/H$ du groupe de Lie G/H est un groupe quotient du groupe fondamental $\pi_1 G$ du groupe de Lie G.*

Démonstration. Soient $\pi: \tilde{G} \to G$ et $\rho: \widetilde{G/H} \to G/H$ des revêtements universels, et soit $\omega: G \to G/H$ une application canonique. Le groupe $\tilde{G}$ étant simplement connexe, l'homomorphisme $\omega \circ \pi : \tilde{G} \to G/H$ est relevable à un homomorphisme $\tilde{\omega}$: $\tilde{G} \to \widetilde{G/H}$. Si $\tilde{H} = \operatorname{Ker} \tilde{\omega}$ et $\tilde{\iota}$: $\tilde{H} \to \tilde{G}$ est une injection, alors, puisque $\omega \circ \pi \circ \tilde{\iota} = \rho \circ \tilde{\omega} \circ \tilde{\iota} = 0$, l'homomorphisme π envoie $\tilde{H}$ dans H et, par suite, induit un homomorphisme π_H: $\tilde{H} \to H$. Ceci est représenté de façon suggestive sur le diagramme commutatif

$$\begin{array}{ccccc} \tilde{H} & \xrightarrow{\tilde{\iota}} & \tilde{G} & \xrightarrow{\tilde{\omega}} & \widetilde{G/H} \\ {\scriptstyle \pi_H}\downarrow & & \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \rho} \\ H & \xrightarrow[\iota]{} & G & \xrightarrow[\omega]{} & G/H \end{array}$$

Il s'avère que si l'homomorphisme ω est un épimorphisme, il induit un épimorphisme du groupe $\pi_1 G = \operatorname{Ker} \pi$ sur le groupe $\pi_1 G/H = \operatorname{Ker} \rho$. En effet, dans ce cas le groupe $\widetilde{G/H}$ sera isomorphe au groupe quotient $\tilde{G}/\tilde{H}$, de sorte que ce groupe quotient sera simplement connexe. Donc, d'après le corollaire de la proposition 6, le sous-groupe $\tilde{H}$ sera connexe et, par suite, l'homomorphisme π_H: $\tilde{H} \to H$ sera un revêtement et, en particulier, un épimorphisme. Si maintenant $\tilde{a} \in \operatorname{Ker} \rho$ et $\tilde{g}$ est un élément de $\tilde{G}$ tel que $\tilde{\omega}(\tilde{g}) = \tilde{a}$, alors l'élément $\pi\tilde{g} \in G$ appartiendra au noyau H de l'épimorphisme ω et sera donc l'image par l'épimorphisme π_H d'un élément $\tilde{h} \in \tilde{H}$, ou, plus exactement, on aura l'égalité $(\iota \circ \pi_H)\tilde{h} = \pi\tilde{g}$. L'élément

$\tilde{g}_1 = (\widetilde{\iota h})^{-1}\tilde{g}$ appartiendra alors au noyau $\operatorname{Ker} \pi$ de l'homomorphisme π, et son image $\tilde{\omega}(\tilde{g}_1)$ par l'homomorphisme $\tilde{\omega}$ sera l'élément considéré $\tilde{a} \in \operatorname{Ker} \rho$.

Donc, pour prouver la proposition 9, il suffit de montrer que $\tilde{\omega}(\tilde{G}) = \widetilde{G/H}$.

Or, il est évident qu'il existe dans G et G/H des bases $\{U_\alpha\}$ et $\{V_\alpha\}$ composées d'ensembles ouverts connexes revêtus sans pli respectivement par les applications π et ρ, telles que pour tout α l'ensemble V_α est l'image $\omega(U_\alpha)$ de U_α par l'épimorphisme ω. Soient $U_{\alpha,\beta}$ les composantes de l'image réciproque $\pi^{-1}(U_\alpha)$ de l'ensemble U_α, et $V_{\alpha,\gamma}$, les composantes de l'image réciproque $\rho^{-1}(V_\alpha)$ de V_α respectivement par les applications π et ρ. Comme chaque ensemble $U_{\alpha,\beta}$ est envoyé par l'homomorphisme $\omega \circ \pi$ dans un ensemble V_α, l'homomorphisme $\tilde{\omega}$ l'envoie dans un ensemble $V_{\alpha,\gamma}$. Donc, si un ensemble $V_{\alpha,\gamma}$ coupe le sous-espace $\tilde{\omega}(\tilde{G})$, il y est alors contenu: $V_{\alpha,\gamma} \subset \tilde{\omega}(\tilde{G})$. Comme les ensembles $V_{\alpha,\gamma}$ forment une base de l'espace $\widetilde{G/H}$, ceci n'est possible que si le sous-espace $\tilde{\omega}(\tilde{G})$ est à la fois ouvert et fermé. Donc, en vertu de la connexité, on a $\tilde{\omega}(\tilde{G}) = \widetilde{G/H}$ (comparer avec la démonstration du lemme 3 de la leçon 8).

Ceci achève la démonstration de la proposition 9. □

Appliquons la proposition 9 à l'épimorphisme $\mathrm{U}(n) \to \mathbb{S}^1$ défini par la formule

$$A \mapsto \frac{\det A}{|\det A|}, \quad A \in \mathrm{U}(n).$$

Comme $\pi_1 \mathbb{S}^1 = \mathbb{Z}$, le groupe $\pi_1 \mathrm{U}(n)$ est épimorphe au groupe $\mathbb{Z}$ d'après cette proposition. Comme, par ailleurs, on a démontré plus haut que le groupe $\mathbb{Z}$ est épimorphe au groupe $\pi_1 \mathrm{U}(n)$, il vient que *le groupe fondamental* $\pi_1 \mathrm{U}(n)$ *du groupe de Lie* $\mathrm{U}(n)$ *est isomorphe au groupe* $\mathbb{Z}$.

Parmi les groupes de Lie de matrices connexes classiques, il nous reste à étudier seulement le groupe $\mathrm{SO}(n)$. C'est ce que nous ferons dans la leçon suivante.

LEÇON 13

Algèbre de Clifford d'une fonctionnelle quadratique.— $\mathbb{Z}_2$-graduation d'une algèbre de Clifford.— Sur le produit tensoriel d'espaces vectoriels et d'algèbres.— Factorisation des algèbres de Clifford en un produit tensoriel gauche.— Base d'une algèbre de Clifford.— Dualité dans une algèbre de Clifford. — Centre d'une algèbre de Clifford.— Le groupe de Lie Spin(n).— Groupe fondamental du groupe SO(n).— Les groupes Spin(n) pour $n \leqslant 4$.— L'homomophisme χ.— Le groupe Spin(6). — Le groupe Spin(5).— Représentations matricielles des algèbres de Clifford.— Représentations matricielles des groupes Spin(n).— Groupes matriciels dans lesquels sont représentés les groupes Spin(n).— Représentations réduites des groupes Spin(n).— Compléments d'algèbre linéaire.

Nous commencerons de loin le calcul du groupe fondamental $\pi_1 SO(n)$ du groupe $SO(n)$.

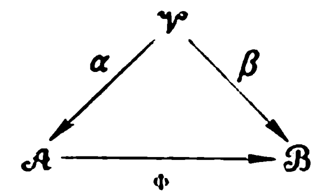
Soit Q une fonctionnelle quadratique (fixée une fois pour toutes) définie sur un espace vectoriel $\mathscr{V}$ (de dimension finie) sur le corps des réels $\mathbb{R}$.

Nous étudierons des couples de la forme $(\mathscr{A}, \alpha)$, où $\mathscr{A}$ est une algèbre unitaire (pas forcément de dimension finie) sur le corps $\mathbb{R}$, et α une application linéaire $\mathscr{V} \to \mathscr{A}$ telle que $\alpha(x)^2 = Q(x)1$ pour tout élément $x \in \mathscr{V}$, où 1 est l'unité de l'algèbre $\mathscr{A}$. (Dans la suite, dans de telles égalités on omettra en général l'unité 1, c'est-à-dire qu'on identifiera les éléments de la forme $\lambda 1 \in \mathscr{A}$ avec les nombres correspondants $\lambda \in \mathbb{R}$.)

Remarque 1. Pour vérifier la condition $\alpha(x)^2 = Q(x)$, il faut se rappeler qu'elle est remplie pour tous les $x \in \mathscr{V}$ si elle l'est pour les éléments d'une base de $\mathscr{V}$.

On appellera *morphisme* d'un couple $(\mathscr{A}, \alpha)$ dans un couple $(\mathscr{B}, \beta)$ un homomorphisme $\Phi: \mathscr{A} \to \mathscr{B}$ tel que l'on ait le dia-

gramme commutatif:



c'est-à-dire tel que $\beta = \Phi \circ \alpha$. Il est évident que les couples $(\mathscr{A}, \alpha)$ et leurs morphismes $(\mathscr{A}, \alpha) \to (\mathscr{B}, \beta)$ forment une catégorie que l'on désignera par CLIFF(Q).

On rappelle qu'un objet A_0 d'une catégorie G est *initial* (ou *universel* dans une autre terminologie) si pour tout objet $A \in$ G il existe un seul morphisme $A_0 \to A$. Il est clair que l'objet initial (s'il existe) est unique à un isomorphisme près.

Proposition 1. *La catégorie* CLIFF(Q) *admet un objet initial.*

D é m o n s t r a t i o n. Soit

$$\mathrm{T}_0(\mathscr{V}) = \mathrm{T}_0^0(\mathscr{V}) \oplus \ldots \oplus \mathrm{T}_0^q(\mathscr{V}) \oplus \ldots,$$

où $\mathrm{T}_0^q(\mathscr{V})$ est l'espace vectoriel des fonctionnelles multilinéaires de la forme $(0, q)$ sur $\mathscr{V}$ (cf. II, 5, page 45). Il est évident que la somme directe $\mathrm{T}_0(\mathscr{V})$ est une algèbre (de dimension infinie) pour la multiplication tensorielle.

Considérons dans l'algèbre $\mathrm{T}_0(\mathscr{V})$ l'idéal $I(Q)$ engendré par les éléments de la forme $x \otimes x - Q(x)$, où $x \in \mathscr{V} = \mathrm{T}_0^1(\mathscr{V})$. Soit $\mathrm{Cl}(Q)$ l'algèbre quotient de l'algèbre $\mathrm{T}_0(\mathscr{V})$ par cet idéal, et soit $\iota\colon \mathscr{V} \to \mathrm{Cl}(Q)$ la restriction à $\mathscr{V} = \mathrm{T}_0^1(\mathscr{V})$ de l'épimorphisme canonique $\mathrm{T}_0(\mathscr{V}) \to \mathrm{Cl}(Q)$. Il s'avère que *le couple* $(\mathrm{Cl}(Q), \iota)$ *est un objet initial de la catégorie* CLIFF(Q).

En effet, par construction $\iota(x)^2 = Q(x)$, de sorte que $(\mathrm{Cl}(Q), \iota) \in$ $\in$ CLIFF(Q). Il est clair par ailleurs que l'algèbre $\mathrm{T}_0(\mathscr{V})$ est isomorphe à l'algèbre $\mathbb{R}\langle x_1, \ldots, x_n\rangle$ des polynômes en n indéterminées non permutables $x_1, \ldots, x_n$ (l'isomorphisme est défini par une base quelconque $e_1, \ldots, e_n$ de l'espace $\mathscr{V}$ et est réalisé par la correspondance $e_{i_1} \otimes \ldots \otimes e_{i_q} \mapsto x_{i_1} \ldots x_{i_q}$). Donc, toute application linéaire $\alpha\colon \mathscr{V} \to \mathscr{A}$, $\mathscr{A}$ étant une algèbre associative et unitaire, se prolonge de façon unique en un homomorphisme $\bar{\alpha}\colon \mathrm{T}_0(\mathscr{V}) \to \mathscr{A}$. Ceci étant, si $\alpha(x)^2 = Q(x)$ pour tout élément $x \in \mathscr{V}$, c'est-à-dire si $(\mathscr{A}, \alpha) \in$ CLIFF(Q), alors $\bar{\alpha} = 0$ sur $I(Q)$ et, par suite, l'homomorphisme $\bar{\alpha}$ induit un homomorphisme $\alpha^{\#}\colon \mathrm{Cl}(Q) \to \mathscr{A}$ tel que visiblement $\alpha^{\#} \circ \iota = \alpha$, c'est-à-dire est un morphisme $(\mathrm{Cl}(Q), \iota) \to (\mathscr{A}, \alpha)$. Comme l'espace vectoriel $\mathscr{V}$ engendre l'algèbre $\mathrm{T}_0(\mathscr{V})$, donc l'espace vectoriel $\iota\mathscr{V}$, l'algèbre

$\mathrm{Cl}(Q)$, ce morphisme est unique. Par conséquent, le couple $(\mathrm{Cl}(Q), \iota)$ est l'objet initial de la catégorie CLIFF(Q). □

Définition 1. L'algèbre $\mathrm{Cl}(Q)$ s'appelle *algèbre de Clifford* de la fonctionnelle quadratique Q.

Par définition, l'algèbre $\mathrm{Cl}(Q)$ est telle que pour chaque objet $(\mathcal{A}, \alpha)$ de la catégorie CLIFF(Q), il existe un seul homomorphisme d'algèbres $\alpha^{\#}\colon \mathrm{Cl}(Q) \to \mathcal{A}$ pour lequel $\alpha^{\#} \circ \iota = \alpha$, et, de plus, elle est totalement caractérisée par cette propriété à un isomorphisme près.

Dans le cas particulier où $\mathcal{V}$ est un espace euclidien et Q une fonctionnelle métrique correspondante (associant à tout vecteur $x \in \mathcal{V}$ le carré $|x|^2$ de son module), l'algèbre $\mathrm{Cl}(Q)$ sera désignée par $\mathrm{Cl}_+(\mathcal{V})$. Si (dans l'hypothèse où $\mathcal{V}$ est encore euclidien) la fonctionnelle Q est définie par la formule $Q(x) = -|x|^2$, alors l'algèbre $\mathrm{Cl}(Q)$ sera notée $\mathrm{Cl}(\mathcal{V})$. Si les algèbres $\mathrm{Cl}_+(\mathcal{V})$ et $\mathrm{Cl}(\mathcal{V})$ sont envisagées simultanément, on les désignera par $\mathrm{Cl}_\varepsilon(\mathcal{V})$, où $\varepsilon = \pm 1$, avec $\varepsilon = +1$ pour $\mathrm{Cl}_+(\mathcal{V})$ et $\varepsilon = -1$ pour $\mathrm{Cl}(\mathcal{V})$.

Si l'on choisit une base orthonormée dans l'espace $\mathcal{V}$, ce qui identifie *ipso facto* $\mathcal{V}$ à l'espace euclidien $\mathbb{R}^n$, on désignera les algèbres $\mathrm{Cl}_+(\mathcal{V})$ et $\mathrm{Cl}(\mathcal{V})$ respectivement par $\mathrm{Cl}_+(n)$ et $\mathrm{Cl}(n)$ (ou par $\mathrm{Cl}_\varepsilon(n)$ si elles sont étudiées simultanément).

Les couples $(\mathcal{V}, Q)$ forment visiblement une catégorie $\mathbf{Q}$ dont les morphismes $(\mathcal{V}, Q) \to (\mathcal{V}_1, Q_1)$ sont des applications linéaires $\varphi\colon \mathcal{V} \to \mathcal{V}_1$ telles que $Q(x) = Q_1(\varphi x)$ pour tout vecteur $x \in \mathcal{V}$. Chacune de ces applications induit de toute évidence un homomorphisme $\mathrm{T}_0(\mathcal{V}) \to \mathrm{T}_0(\mathcal{V}_1)$ envoyant l'idéal $I(Q)$ dans l'idéal $I(Q_1)$ et, par suite, induit un homomorphisme

$$\mathrm{Cl}\varphi\colon \mathrm{Cl}(Q) \to \mathrm{Cl}(Q_1).$$

Il est évident que les correspondances $(\mathcal{V}, Q) \mapsto \mathrm{Cl}(Q)$, $\varphi \mapsto \mathrm{Cl}\,\varphi$ forment un foncteur Cl de la catégorie $\mathbf{Q}$ dans la catégorie ALG_0-ASS des algèbres unitaires associatives.

A noter que $\mathrm{Cl}\,\varphi$ n'est autre que l'application $(\iota_1 \circ \varphi)^{\#}$, où $\iota_1\colon \mathcal{V}_1 \to \mathrm{Cl}(Q_1)$ est une application canonique. Par abus de langage, on désignera l'homomorphisme $\mathrm{Cl}\,\varphi$ par $\varphi^{\#}$.

Comme $\iota(x)^2 = Q(x)$, de $\iota(x) = \iota(y)$ il s'ensuit que $Q(x) = Q(y)$. Donc, la correspondance $\iota(x) \to Q(x)$ définit sans ambiguïté une fonctionnelle quadratique sur le sous-espace $\iota\mathcal{V} \subset \mathrm{Cl}(Q)$. (En fait, on montrera plus bas que de $\iota(x) = \iota(y)$, il résulte que $x = y$, de sorte que toutes ces précautions sont pratiquement superflues.)

Pour alléger les formules, on désignera la fonctionnelle $\iota(x) \mapsto Q(x)$ par Q et l'élément $\iota(x)$ par x. Ainsi, pour tout élément $x \in \iota\mathcal{V}$, on aura l'égalité

$$x^2 = Q(x). \tag{1}$$

En particulier, $(x+y)^2 = Q(x+y)$, c'est-à-dire que $x^2 + xy + yx + y^2 = Q(x) + 2Q(x, y) + Q(y)$ et, par suite,

(2) $$xy + yx = 2Q(x, y)$$

pour tous éléments $x, y \in \iota\mathcal{V}$.

Supposons que l'application $-\iota: \mathcal{V} \to \mathrm{Cl}(Q)$ est définie par la formule $(-\iota)(x) = -\iota(x)$. Il est évident que le couple $(\mathrm{Cl}(Q), -\iota)$ appartient à la catégorie CLIFF(Q). Donc est défini l'homomorphisme $\alpha = (-\iota)^{\#}$, c'est-à-dire un homomorphisme $\alpha: \mathrm{Cl}(Q) \to \mathrm{Cl}(Q)$ tel que $\alpha x = -x$ si $x \in \iota\mathcal{V}$. Il est évident que $\alpha^2 = \mathrm{id}$, c'est-à-dire que l'homomorphisme α est un automorphisme involutif. Pour tout élément $u \in \mathrm{Cl}(Q)$, l'élément αu sera désigné par u^*.

Pour étudier l'automorphisme $u \to u^*$, on se servira du lemme suivant d'algèbre linéaire:

Lemme 1. *Si un opérateur linéaire $A: \mathcal{W} \to \mathcal{W}$ agissant dans un espace réel ou complexe $\mathcal{W}$ est involutif, c'est-à-dire que $A^2 = E$, alors ses valeurs propres sont égales à ± 1 et il est diagonalisable, c'est-à-dire que l'espace $\mathcal{W}$ est la somme directe*

(3) $$\mathcal{W} = \mathcal{W}_+ \oplus \mathcal{W}_-$$

des espaces propres $\mathcal{W}_+$ et $\mathcal{W}_-$ associés respectivement aux valeurs propres $+1$ et -1.

Démonstration. Supposons que le corps de base est le corps $\mathbb{C}$ des nombres complexes. L'opérateur A est alors justiciable du théorème de réduction à la forme normale de Jordan (cf. II, 16), c'est-à-dire qu'il est somme directe d'opérateurs de la forme $\lambda E + C$, où $\lambda \in \mathbb{C}$ et C est soit un opérateur nul, soit un opérateur cyclique. Ceci étant, si l'opérateur A est involutif, il en est de même de chacun des opérateurs de la somme $\lambda E + C$. Or $(\lambda E + C)^2 = \lambda^2 E + 2\lambda C + C^2$, donc l'égalité $(\lambda E + C)^2 = E$ n'est possible que si $C = 0$. Ce qui prouve que l'opérateur A est diagonalisable. Ceci achève la démonstration du lemme 1 dans le cas où le corps de base est $\mathbb{C}$, puisque les éléments diagonaux d'une matrice diagonale involutive sont de toute évidence égaux à ± 1.

Si l'espace vectoriel $\mathcal{W}$ est réel, on passe à son complexifié $\mathcal{W}^{\mathbb{C}}$ (cf. II, 17). L'opérateur complexifié $A^{\mathbb{C}}$ étant manifestement encore involutif, d'après ce qui a été démontré on a

$$\mathcal{W}^{\mathbb{C}} = \mathcal{W}_+^{\mathbb{C}} \oplus \mathcal{W}_-^{\mathbb{C}}.$$

En se limitant dans cette décomposition aux seuls vecteurs réels, on obtient aisément la décomposition (3). □

En vertu de ce lemme appliqué à l'automorphisme $u \mapsto u^*$, l'espace vectoriel $\mathrm{Cl}(Q)$ se décompose en la somme directe

$$\mathrm{Cl}(Q) = \mathrm{Cl}^0(Q) \oplus \mathrm{Cl}^1(Q)$$

des sous-espaces $\mathcal{C}l^0(Q)$ et $\mathcal{C}l^1(Q)$ composés, le premier, des éléments tels que $u^* = u$, le second, des éléments, tels que $u^* = -u$.

Les éléments de $\mathcal{C}l^0(Q)$ s'appellent *éléments pairs* et ceux de $\mathcal{C}l^1(Q)$, *éléments impairs* de l'algèbre de Clifford $\mathcal{C}l(Q)$.

Il est évident que le produit de deux éléments pairs ou impairs est pair, et celui d'un élément pair par un impair est impair, c'est-à-dire que

$$\mathcal{C}l^i(Q)\cdot\mathcal{C}l^j(Q)\subset\mathcal{C}l^{i+j}(Q)$$

pour tous $i, j = 0, 1$ (il s'agit d'une sommation modulo 2).

On voit, en particulier, que *le sous-espace* $\mathcal{C}l^0(Q)$ *est une sous-algèbre de l'algèbre* $\mathcal{C}l(Q)$.

La structure algébrique obtenue mérite un nom spécial.

Définition 2. On dit qu'une algèbre $\mathcal{A}$ est *$\mathbb{Z}_2$-graduée* si $\mathcal{A} = \mathcal{A}^0 \oplus \mathcal{A}^1$ et de plus

$$\mathcal{A}^i\cdot\mathcal{A}^j\subset\mathcal{A}^{i+j \bmod 2}$$

pour tous $i, j = 0, 1$. On appelle *morphisme* d'algèbres $\mathbb{Z}_2$-graduées un homomorphisme $\varphi\colon \mathcal{A}\to\mathcal{B}$ tel que $\varphi(\mathcal{A}^i)\subset\mathcal{B}^i$ pour tout $i = 0, 1$. Il est évident que les algèbres $\mathbb{Z}_2$-graduées et leurs morphismes forment une catégorie. Désignons cette catégorie par $\mathbb{Z}_2$-ALG et sa sous-catégorie formée par les algèbres associatives unitaires, par $\mathbb{Z}_2$-ALG$_0$-ASS.

D'après ce qui précède, le foncteur $\mathcal{C}l$ peut être traité comme un foncteur de la catégorie $\mathbf{Q}$ dans la catégorie $\mathbb{Z}_2$-ALG$_0$-ASS. Ceci étant, il est immédiat de voir que si pour un couple $(\mathcal{A}, \alpha) \in \in$ CLIFF(Q), l'algèbre $\mathcal{A}$ est $\mathbb{Z}_2$-graduée et l'application $\alpha\colon \mathcal{V}\to\mathcal{A}$ est une application dans $\mathcal{A}^1$, alors l'homomorphisme correspondant $\alpha^\#\colon \mathcal{C}l(Q)\to\mathcal{A}$ est un morphisme d'algèbres $\mathbb{Z}_2$-graduées.

A la leçon 5, nous avons introduit la notion de produit tensoriel d'algèbres et d'espaces vectoriels. On rappelle que les éléments de l'espace vectoriel $\mathcal{A}\otimes\mathcal{B}$ sont les combinaisons linéaires des symboles $a\otimes b$, où $a\in\mathcal{A}$ et $b\in\mathcal{B}$, et de plus

$$(a_1+a_2)\otimes b=a_1\otimes b+a_2\otimes b,$$

$$a\otimes(b_1+b_2)=a\otimes b_1+a\otimes b_2$$

pour tous éléments $a_1, a_2, a\in\mathcal{A}$ et $b, b_1, b_2\in\mathcal{B}$. Si $e_1, \ldots, e_n$ est une base de $\mathcal{A}$ et $f_1, \ldots, f_m$, une base de $\mathcal{B}$, les éléments $e_i\otimes\otimes f_j$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, forment une base dans $\mathcal{A}\otimes\mathcal{B}$ (de sorte que, en particulier, $\dim(\mathcal{A}\otimes\mathcal{B}) = \dim\mathcal{A}\cdot\dim\mathcal{B}$).

De cette description de l'espace vectoriel $\mathcal{A}\otimes\mathcal{B}$ il résulte immédiatement que pour tous espaces vectoriels $\mathcal{A}, \mathcal{B}, \mathcal{C}$ on a les

isomophismes canoniques

(4) $\mathcal{A} \otimes \mathcal{B} \approx \mathcal{B} \otimes \mathcal{A}$ (commutativité),

(5) $(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C} \approx \mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C})$ (associativité),

(6) $(\mathcal{A} \oplus \mathcal{B}) \otimes \mathcal{C} \approx (\mathcal{A} \otimes \mathcal{C}) \oplus (\mathcal{B} \otimes \mathcal{C})$ (distributivité).

Si $\mathcal{A}$ et $\mathcal{B}$ sont des algèbres, on munit $\mathcal{A} \otimes \mathcal{B}$ d'une multiplication telle que

$$(a \otimes b)(a_1 \otimes b_1) = aa_1 \otimes bb_1 \tag{7}$$

pour tous éléments $a, a_1 \in \mathcal{A}, b, b_1 \in \mathcal{B}$. L'espace $\mathcal{A} \otimes \mathcal{B}$ est une algèbre pour cette multiplication et les isomorphismes (4), (5) et (6), des isomorphismes d'algèbres. (Si $\mathcal{A}$ et $\mathcal{B}$ sont des algèbres, alors la somme directe $\mathcal{A} \oplus \mathcal{B}$ est munie d'une multiplication « composante par composante », c'est-à-dire que $(a, b)(a_1, b_1) = (aa_1, bb_1)$.) Si les algèbres $\mathcal{A}$ et $\mathcal{B}$ sont associatives et unitaires, il en est de même de l'algèbre $\mathcal{A} \otimes \mathcal{B}$ (l'unité étant $1 \otimes 1$).

Soient $\mathcal{A} = \mathcal{A}^0 \oplus \mathcal{A}^1$ et $\mathcal{B} = \mathcal{B}^0 \oplus \mathcal{B}^1$ des algèbres $\mathbb{Z}_2$-graduées. En posant

$$(\mathcal{A} \otimes \mathcal{B})^0 = (\mathcal{A}^0 \otimes \mathcal{B}^0) \oplus (\mathcal{A}^1 \otimes \mathcal{B}^1),$$
$$(\mathcal{A} \otimes \mathcal{B})^1 = (\mathcal{A}^0 \otimes \mathcal{B}^1) \oplus (\mathcal{A}^1 \otimes \mathcal{B}^0),$$

il vient immédiatement en vertu de (4) et (6) que

$$\mathcal{A} \otimes \mathcal{B} = (\mathcal{A} \otimes \mathcal{B})^0 \oplus (\mathcal{A} \otimes \mathcal{B})^1.$$

Bien que l'algèbre $\mathcal{A} \otimes \mathcal{B}$ soit une algèbre $\mathbb{Z}_2$-graduée pour cette décomposition, il apparaît plus commode de munir $\mathcal{A} \otimes \mathcal{B}$ d'une autre multiplication telle que

$$(a \otimes b)(a_1 \otimes b_1) = (-1)^{ij}(aa_1 \otimes bb_1), \tag{8}$$

où $b \in \mathcal{B}^i$ et $a_1 \in \mathcal{A}^j$. Il est évident que pour cette multiplication l'espace vectoriel $\mathcal{A} \otimes \mathcal{B}$ sera une algèbre $\mathbb{Z}_2$-graduée unitaire et associative si les algèbres $\mathcal{A}$ et $\mathcal{B}$ sont unitaires et associatives.

L'algèbre $\mathcal{A} \otimes \mathcal{B}$ munie de la multiplication (8) s'appelle *produit tensoriel gauche* des algèbres $\mathbb{Z}_2$-graduées $\mathcal{A}$ et $\mathcal{B}$. Quand on aura à distinguer ce produit du produit tensoriel ordinaire (quoique $\mathbb{Z}_2$-gradué) $\mathcal{A} \otimes \mathcal{B}$, on le désignera par $\mathcal{A} \hat{\otimes} \mathcal{B}$.

Ces trois produits tensoriels (pour les espaces vectoriels, les algèbres et les algèbres $\mathbb{Z}_2$-graduées) possèdent la propriété de fonctorialité, c'est-à-dire pour tous morphismes $\varphi : \mathcal{A} \to \mathcal{A}_1$, $\psi : \mathcal{B} \to \mathcal{B}_1$ est défini un morphisme $\varphi \otimes \psi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A}_1 \otimes \mathcal{B}_1$ satisfaisant les identités fonctorielles habituelles. Ce morphisme est déterminé de façon unique par la relation

$$(\varphi \otimes \psi)(a \otimes b) = \varphi(a) \otimes \psi(b),$$

qui doit être vérifiée pour tous éléments $a \in \mathcal{A}$, $b \in \mathcal{B}$.

Par ailleurs, pour tous espaces vectoriels $\mathcal{A}$ et $\mathcal{B}$ et toute algèbre $\mathcal{C}$, on peut à des applications linéaires quelconques $\varphi\colon \mathcal{A} \to \mathcal{C}$ et $\psi\colon \mathcal{B} \to \mathcal{C}$ associer l'application $\varphi \otimes \psi\colon \mathcal{A} \otimes \mathcal{B} \to \mathcal{C}$ qui est définie de façon unique par la relation

$$(\varphi \otimes \psi)(a \otimes b) = \varphi(a)\psi(b), \quad a \in \mathcal{A}, \quad b \in \mathcal{B}.$$

Si $\mathcal{A}$ et $\mathcal{B}$ sont des algèbres et $\varphi\colon \mathcal{A} \to \mathcal{C}$ et $\psi\colon \mathcal{B} \to \mathcal{C}$ des homomorphismes qui commutent, c'est-à-dire que pour tous éléments $a \in \mathcal{A}$ et $b \in \mathcal{B}$ on a $\varphi a \cdot \psi b = \psi b \cdot \varphi a$, alors l'application $\varphi \otimes \otimes \psi\colon \mathcal{A} \otimes \mathcal{B} \to \mathcal{C}$ sera aussi un homomorphisme. De façon analogue, pour toutes algèbres $\mathbb{Z}_2$-graduées $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ et tous morphismes *anticommutatifs* $\varphi\colon \mathcal{A} \to \mathcal{C}$, $\psi\colon \mathcal{B} \to \mathcal{C}$ (c'est-à-dire tels que $\varphi a \cdot \psi b = (-1)^{ij} \psi b \cdot \varphi a$, $a \in \mathcal{A}^i$, $b \in \mathcal{B}^j$), l'application $\varphi \otimes \psi\colon \mathcal{A} \otimes \otimes \mathcal{B} \to \mathcal{C}$ sera morphisme du produit tensoriel gauche $\mathcal{A} \hat{\otimes} \mathcal{B}$ dans l'algèbre $\mathcal{C}$.

Pour tout couple d'objets $(\mathcal{V}_1, Q_1)$ et $(\mathcal{V}_2, Q_2)$ de la catégorie $\mathbf{Q}$, la formule

$$Q(x_1 + x_2) = Q_1(x_1) + Q_2(x_2), \quad x_1 \in \mathcal{V}_1, \quad x_2 \in \mathcal{V}_2,$$

définit sur la somme directe $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$ une fonctionnelle quadratique Q appelée *somme directe des fonctionnelles quadratiques* Q_1 et Q_2 (et désignée généralement par $Q_1 \oplus Q_2$).

Traduit dans ce langage, le théorème de Lagrange (cf. II, 11) dit que toute fonctionnelle quadratique est somme directe de fonctionnelles sur des espaces à une dimension.

Proposition 2. *Pour tout couple de fonctionnelles quadratiques* Q_1 *et* Q_2, *on a l'isomorphisme canonique*

$$\mathrm{Cl}(Q_1 \oplus Q_2) \approx \mathrm{Cl}(Q_1) \hat{\otimes} \mathrm{Cl}(Q_2).$$

D é m o n s t r a t i o n. Posons $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$ et définissons une application linéaire

$$\alpha\colon \mathcal{V} \to \mathrm{Cl}(Q_1) \hat{\otimes} \mathrm{Cl}(Q_2)$$

par la formule

$$\alpha(x_1 + x_2) = x_1 \otimes 1 + 1 \otimes x_2, \quad x_1 \in \mathcal{V}_1, \quad x_2 \in \mathcal{V}_2, \tag{9}$$

où comme toujours $x_1 = \iota x_1$, $x_2 = \iota x_2$. Comme

$$(x_1 \otimes 1)(1 \otimes x_2) = x_1 \otimes x_2$$

et

$$(1 \otimes x_2)(x_1 \otimes 1) = -(x_1 \otimes x_2),$$

il vient

$$\alpha(x)^2 = \alpha(x_1 + x_2)^2 = (x_1 \otimes 1 + 1 \otimes x_2)^2 =$$
$$= x_1^2 \otimes 1 + 1 \otimes x_2^2 =$$
$$= Q_1(x_1) + Q_2(x_2) =$$
$$= Q(x)$$

pour tout vecteur $x = x_1 + x_2 \in \mathscr{V}$, où $x_1 \in \mathscr{V}_1$, $x_2 \in \mathscr{V}_2$, $Q = Q_1 \oplus Q_2$, et, par suite, $(\mathrm{Cl}(Q_1) \hat{\otimes} \mathrm{Cl}(Q_2), \alpha) \in \mathrm{CLIFF}(Q)$. Montrons que *le morphisme correspondant*

$$\alpha^{\#}: \mathrm{Cl}(Q) \to \mathrm{Cl}(Q_1) \hat{\otimes} \mathrm{Cl}(Q_2)$$

d'algèbres $\mathbb{Z}_2$-graduées est un isomorphisme.

Considérons à cet effet les injections canoniques $\sigma_1: \mathscr{V}_1 \to \mathscr{V}$ et $\sigma_2: \mathscr{V}_2 \to \mathscr{V}$. Il est clair que ce sont des morphismes respectivement des couples $(\mathscr{V}_1, Q_1)$ et $(\mathscr{V}_2, Q_2)$ dans le couple $(\mathscr{V}, Q)$. Donc sont définis les homomorphismes $\sigma_1^{\#}: \mathrm{Cl}(Q_1) \to \mathrm{Cl}(Q)$ et $\sigma_2^{\#}: \mathrm{Cl}(Q_2) \to \mathrm{Cl}(Q)$. Comme pour tous vecteurs $x_1 \in \mathscr{V}_1$, $x_2 \in \mathscr{V}_2$, les vecteurs $\sigma_1 x_1$, $\sigma_2 x_2 \in \mathscr{V}$, sont par définition Q-orthogonaux, on déduit d'après la formule (2) que

$$\sigma_1^{\#} x_1 \cdot \sigma_2^{\#} x_2 = -\sigma_2^{\#} x_2 \cdot \sigma_1^{\#} x_1.$$

Comme tout élément pair (resp. impair) de l'algèbre $\mathrm{Cl}(Q_i)$, $i = 1, 2$, est la somme de produits d'un nombre pair (resp. impair) d'éléments $x_i \in \iota\mathscr{V}_i$, il s'ensuit immédiatement que les applications $\sigma_1^{\#}$ et $\sigma_2^{\#}$ anticommutent et, par suite, l'application $\sigma_1^{\#} \otimes \sigma_2^{\#}: \mathrm{Cl}(Q_1) \hat{\otimes} \mathrm{Cl}(Q_2) \to \mathrm{Cl}(Q)$ est un morphisme de l'algèbre $\mathbb{Z}_2$-graduée $\mathrm{Cl}(Q_1) \hat{\otimes} \mathrm{Cl}(Q_2)$ dans l'algèbre $\mathbb{Z}_2$-graduée $\mathrm{Cl}(Q)$.

Sous la forme complète (c'est-à-dire avec σ_1 et σ_2) la formule (9) s'écrit

$$\alpha(\sigma_1 x_1 + \sigma_2 x_2) = x_1 \otimes 1 + 1 \otimes x_2.$$

Comme, par définition, $\sigma_1^{\#} x_1 = \iota\sigma_1 x_1$, $\sigma_2^{\#} x_2 = \iota\sigma_2 x_2$ et $\alpha^{\#} \circ \iota = \alpha$, on en déduit que

$$\alpha^{\#}(\sigma_1^{\#} x_1 + \sigma_2^{\#} x_2) = x_1 \otimes 1 + 1 \otimes x_2$$

pour tous éléments $x_1 \in \iota\mathscr{V}_1$, $x_2 \in \iota\mathscr{V}_2$. Ici $\sigma_1^{\#} x_1 + \sigma_2^{\#} x_2$ est un élément arbitraire x de $\iota\mathscr{V}$, donc

$$((\sigma_1^{\#} \otimes \sigma_2^{\#}) \circ \alpha^{\#})x = (\sigma_1^{\#} \otimes \sigma_2^{\#})(x_1 \otimes 1 + 1 \otimes x_2) =$$
$$= \sigma_1^{\#} x_1 \cdot 1 + 1 \cdot \sigma_2^{\#} x_2 =$$
$$= \sigma_1^{\#} x_1 + \sigma_2^{\#} x_2 = x,$$

de sorte que $(\sigma_1^\# \otimes \sigma_2^\#) \circ \alpha^\# = \mathrm{id}$ sur le sous-espace $\iota\mathcal{V}$. Comme l'application $(\sigma_1^\# \otimes \sigma_2^\#) \circ \alpha^\#$ est un homomorphisme d'algèbres et que le sous-espace $\iota\mathcal{V}$ engendre l'algèbre $\mathrm{Cl}(Q)$, on en déduit que $(\sigma_1^\# \otimes \sigma_2^\#) \circ \alpha^\# = \mathrm{id}$ sur l'algèbre $\mathrm{Cl}(Q)$ tout entière.

De façon analogue

$$(\alpha^\# \circ (\sigma_1^\# \otimes \sigma_2^\#))(x_1 \otimes 1) = \alpha^\#(\sigma_1^\# x_1 \cdot 1) = x_1 \otimes 1$$

et

$$(\alpha^\# \circ (\sigma_1^\# \otimes \sigma_2^\#))(1 \otimes x_2) = \alpha^\#(1 \cdot \sigma_2^\# x_2) = 1 \otimes x_2$$

pour tous éléments $x_1 \in \iota\mathcal{V}_1$, $x_2 \in \iota\mathcal{V}_2$, et, par suite, $\alpha^\# \circ (\sigma_1^\# \otimes \sigma_2^\#) = \mathrm{id}$, puisque les éléments $x_1 \otimes 1$, $x_1 \in \iota\mathcal{V}_1$, et $1 \otimes x_2$, $x_2 \in \iota\mathcal{V}_2$, engendrent l'algèbre $\mathrm{Cl}(Q_1) \hat{\otimes} \mathrm{Cl}(Q_2)$.

Donc, les morphismes $\alpha^\#$ et $\sigma_1^\# \otimes \sigma_2^\#$ sont des isomorphismes réciproques l'un de l'autre. □

Soit $n = \dim \mathcal{V}$. Pour $n = 1$ seules (à un isomorphisme près) trois fonctionnelles quadratiques Q_{+1}, Q_{-1} et Q_0 prennent respectivement les valeurs $+1$, -1 et 0 sur un vecteur $e \in \mathcal{V}$ (base de l'espace à une dimension $\mathcal{V}$). Comme pour $n = 1$ l'algèbre $\mathrm{T}_0(\mathcal{V})$ est isomorphe à l'algèbre $\mathbb{R}[e]$ des polynômes (ordinaires) de e, il s'ensuit de là que l'algèbre $\mathrm{Cl}(Q_\varepsilon)$, $\varepsilon = \pm 1, 0$, se déduit de l'algèbre $\mathbb{R}[e]$ en posant $e^2 = \varepsilon$, c'est-à-dire soit est l'algèbre $\mathbb{C}$ des nombres complexes, soit l'algèbre $\mathbb{D}$ des *nombres doubles* (de la forme $a + be$, où $a, b \in \mathbb{R}$ et $e^2 = 1$), soit l'algèbre des *nombres duaux* (de la forme $a + be$, où $a, b \in \mathbb{R}$ et $e^2 = 0$).

Théorème 1. *Pour toute fonctionnelle quadratique Q sur un espace vectoriel $\mathcal{V}$ à n dimensions, l'algèbre de Clifford $\mathrm{Cl}(Q)$ est isomorphe au produit tensoriel gauche de p algèbres de nombres doubles, $r - p$ algèbres de nombres complexes et $n - r$ algèbres de nombres duaux, où r est le rang de le fonctionnelle Q et p son indice d'inertie positif. En particulier,*

$$\mathrm{Cl}(n) = \underbrace{\mathbb{C} \hat{\otimes} \ldots \hat{\otimes} \mathbb{C}}_{n \text{ fois}}, \quad \mathrm{Cl}_+(n) = \underbrace{\mathbb{D} \hat{\otimes} \ldots \hat{\otimes} \mathbb{D}}_{n \text{ fois}}.$$

D é m o n s t r a t i o n. Résulte par une récurrence immédiate de la proposition 2 et du théorème de Lagrange. □

Etant donné que les algèbres voient leurs dimensions se multiplier quand on fait leur produit tensoriel, il s'ensuit du théorème 1 que $\dim \mathrm{Cl}(Q) = 2^n$.

Bien plus, si $e_1, \ldots, e_n$ est une base Q-orthogonale de l'espace $\mathcal{V}$, alors comme la base du i-ième facteur est composée des éléments 1 et e_i, la base de l'algèbre $\mathrm{Cl}(Q)$ sera composée de 2^n éléments de

la forme $x_{k_1}^{(1)}x_{k_2}^{(2)}\ldots x_{k_n}^{(n)}$, où $k_1, k_2, \ldots, k_n = 0, 1$, et $x_0^{(i)} = 1$, $x_1^{(i)} = e_i$ pour tout $i = 1, \ldots, n$.

Introduisons le sous-ensemble I de la portion $[n] = \{1, 2, \ldots, n\}$ de la série des entiers naturels, composé des indices i tels que $k_i = 1$, et appelons e_I l'élément $x_{k_1}^{(1)}x_{k_2}^{(2)}\ldots x_{k_n}^{(n)}$. Donc, si $I = \{i_1 < \ldots < i_m\}$, alors $e_I = e_{i_1}\ldots e_{i_m}$.

On désignera m par $|I|$.

En particulier, pour $m = 1$ on obtient des éléments $e_{\{1\}} = e_1, \ldots, e_{\{n\}} = e_n$. *Ces éléments sont donc linéairement indépendants* et, par suite, l'application $\iota\colon \mathcal{V} \to \mathrm{Cl}(Q)$ (qui, rappelons-le, envoie les vecteurs de base $e_1, \ldots e_n$ dans les éléments $e_1, \ldots, e_n$) est, comme on l'a déjà affirmé plus haut, un monomorphisme. On identifiera en principe chaque vecteur $x \in \mathcal{V}$ à l'élément associé $x \in \iota\mathcal{V} \in \mathrm{Cl}(Q)$. En vertu de cette convention, pour tout objet $(\mathcal{A}, \alpha) \in \mathrm{CLIFF}(Q)$ le morphisme $\alpha^{\#}\colon \mathrm{Cl}(Q) \to \mathcal{A}$ n'est autre que le prolongement de l'application α de $\mathcal{V}$ à $\mathrm{Cl}(Q)$.

Pour $m = 0$, c'est-à-dire pour $I = \varnothing$, l'élément $e_\varnothing$ (désigné aussi par e_0) est l'unité 1 de l'algèbre $\mathrm{Cl}(Q)$.

Les vecteurs e_i et e_j étant par hypothèse Q-orthogonaux, il vient d'après la formule (2)

$$e_ie_j + e_je_i = 0, \quad i \neq j. \tag{10}$$

Par ailleurs, d'après la formule (1)

$$e_i^2 = \varepsilon_i, \quad \text{où} \quad \varepsilon_i = Q(e_i). \tag{11}$$

Les relations (10) et (11) nous permettent d'écrire immédiatement le produit d'éléments de base quelconques e_I et e_J. Par exemple, pour $n = r$ et $p = 0$, c'est-à-dire dans l'algèbre $\mathrm{Cl}_+(n)$, on a la formule

$$e_Ie_J = (-1)^{\tau_+(I, J)}\, e_{I\triangle J}, \tag{12}$$

où $I \triangle J = (I \cup J) \setminus (I \cap J)$ est la différence symétrique des ensembles I et J, et $\tau_+(I, J)$, le nombre des couples $(i, j) \in I \times J$ tels que $i > j$.

De façon analogue, dans l'algèbre $\mathrm{Cl}(n)$

$$e_Ie_J = (-1)^{\tau(I, J)} e_{I\triangle J}, \tag{13}$$

où $\tau(I, J)$ est le nombre des couples $(i, j) \in I \times J$ tels que $i \geqslant j$.

On dit qu'une application linéaire $a \mapsto \bar{a}$ d'une algèbre $\mathcal{A}$ pas forcément associative dans elle-même est un *antiautomorphisme involutif* si $\bar{\bar{a}} = a$ et $\overline{ab} = \bar{b}\bar{a}$ pour tous éléments $a, b \in \mathcal{A}$. Exemple: l'application de l'algèbre tensorielle $\mathsf{T}_0(\mathcal{V})$ dans elle-même, définie sur les générateurs $x_1 \otimes \ldots \otimes x_p$ par la formule

$$\overline{x_1 \otimes \ldots \otimes x_p} = x_p \otimes \ldots \otimes x_1.$$

Comme tous les éléments de la forme $x \otimes x - Q(x)$ sont invariants par cet antiautomorphisme, il vient que $\overline{I(Q)} = I(Q)$ et, par suite, l'antiautomorphisme considéré induit un antiautomorphisme involutif de l'algèbre $\mathrm{Cl}(Q)$.

Définition 3. L'antiautomorphisme involutif $a \mapsto \bar{a}$ de l'algèbre $\mathrm{Cl}(Q)$ s'appelle *conjugaison*.

Il est immédiat de voir que la conjugaison agit sur les éléments de la base e_I, $I = \{i_1 < \ldots < i_m\} \subset [n]$, à l'aide de la formule

$$\bar{e}_I = (-1)^{\frac{m(m-1)}{2}} e_I.$$

Donc, $\bar{e}_I = e_I$ pour $m = 4p$, $4p + 1$ et $\bar{e}_I = -e_I$ pour $m = 4p + 2$, $4p + 3$.

On rappelle que le *centre* d'une algèbre associative est la sous-algèbre composée de tous les éléments permutables à chaque élément de cette algèbre.

Calculons le centre de l'algèbre de Clifford $\mathrm{Cl}(Q)$ dans le cas où la fonctionnelle Q est définie positive ou négative, c'est-à-dire pour les algèbres $\mathrm{Cl}_\varepsilon(n)$.

Proposition 3. *Si n est pair, le centre de l'algèbre $\mathrm{Cl}_\varepsilon(n)$ est à une dimension (et est composé des seuls éléments de* $\mathbb{R}$), *si n est impair, le centre de l'algèbre $\mathrm{Cl}_\varepsilon(n)$ est à deux dimensions et est engendré par les éléments* 1 *et* $e_{[n]} = e_1 e_2 \ldots e_n$.

D é m o n s t r a t i o n. Comme les éléments $e_1, \ldots, e_n$ engendrent l'algèbre $\mathrm{Cl}_\varepsilon(n)$, un élément $x \in \mathrm{Cl}_\varepsilon(n)$ appartient au centre de cette algèbre si et seulement si $xe_i = e_i x$, c'est-à-dire que $e_i x e_i = \varepsilon x$, pour tout $i = 1, \ldots, n$. Or, il est clair que si $i \in I = \{i_1 < \ldots < i_m\}$ et $i = i_t$, alors

$$e_i e_I = (-1)^{t-1} \varepsilon e_{I \setminus \{i\}} \quad \text{et} \quad e_I e_i = (-1)^{m-t} \varepsilon e_{I \setminus \{i\}},$$

et si $i \notin I$ et $i_{t-1} < i < i_t$, alors

$$e_i e_I = (-1)^{t-1} e_{I \cup \{i\}} \quad \text{et} \quad e_I e_i = (-1)^{m-t+1} e_{I \cup \{i\}}.$$

Donc, si $x = \sum_I x_I e_I$, alors

$$e_i x e_i = \sum_{i \in I} (-1)^{m+1} \varepsilon x_I e_I + \sum_{i \notin I} (-1)^m \varepsilon x_I e_I, \text{ où } m = |I|.$$

Par conséquent, $e_i x e_i = \varepsilon x_i$ si et seulement si $x_I = (-1)^{m+1} x_I$ pour $i \in I$ et $x_I = (-1)^m x_I$ pour $i \notin I$. Comme pour tout $I \neq \varnothing$, $[n]$, il existe des éléments $i \in I$ et des éléments $i \notin I$, on en déduit que si $e_i x e_i = \varepsilon x$ pour tous les $i = 1, \ldots, n$, alors $(-1)^{m+1} x_I = (-1)^m x_I$ et, par suite, $x_I = 0$. Par ailleurs, $x_{[n]} = (-1)^{n+1} x_{[n]}$

et $x_{[n]} = 0$ si n est pair. Donc, $x = x_{\mathcal{S}}e_{\mathcal{S}}$ si n est pair et $x = x_{\mathcal{S}}e_{\mathcal{S}} + x_{[n]}e_{[n]}$ si n est impair. □

Corollaire. *Pour tout n et tout ε les seuls éléments pairs du centre de l'algèbre* $\mathrm{Cl}_\varepsilon(n)$ *sont les nombres de* $\mathbb{R}$. □

Comme $x^2 = \varepsilon|x|^2$ pour tout élément $x \in \iota\mathbb{R}^n = \mathbb{R}^n$, tous les éléments non nuls de $\mathbb{R}^n$ sont inversibles dans l'algèbre $\mathrm{Cl}_\varepsilon(n)$. En particulier, sont inversibles tous les éléments de la sphère unité S^{n-1} de l'espace $\mathbb{R}^n$, et, de plus, si $x \in S^{n-1}$, alors $x^{-1} = \varepsilon x = \varepsilon\bar{x}$.

Définition 4. Le sous-groupe $\mathrm{pin}_\varepsilon(n)$ du groupe multiplicatif de tous les éléments inversibles de l'algèbre $\mathrm{Cl}_\varepsilon(n)$, engendré par les éléments de S^{n-1} s'appelle *groupe de Clifford* de degré n et d'indice ε. Pour $\varepsilon = +1$, ce groupe est noté $\mathrm{pin}_+(n)$ et pour $\varepsilon = -1$, $\mathrm{pin}(n)$.

Le sous-groupe $\mathrm{pin}_\varepsilon(n)$ composé des éléments pairs est noté $\mathrm{Spin}_\varepsilon(n)$ ($\mathrm{Spin}_+(n)$ pour $\varepsilon = +1$ et $\mathrm{Spin}(n)$ pour $\varepsilon = -1$) et appelé *groupe des spineurs* de degré n et d'indice ε.

Il est clair que les groupes $\mathrm{pin}_\varepsilon(n)$ et $\mathrm{Spin}_\varepsilon(n)$ sont fermés dans le groupe de Lie de tous les éléments inversibles de l'algèbre $\mathrm{Cl}_\varepsilon(n)$. Ces groupes sont donc des groupes de Lie.

Par définition, tout élément u du groupe $\mathrm{pin}_\varepsilon(n)$ se représente (en général de façon non unique) sous forme d'un produit $x_1 \ldots x_m$, où $x_1, \ldots, x_m \in S^{n-1}$, et cet élément appartient à $\mathrm{Spin}_\varepsilon(n)$ si et seulement si m est pair. Ceci étant, comme les applications $u \mapsto u^{-1}$ et $u \mapsto \bar{u}$ sont des antiautomorphismes et $x^{-1} = \varepsilon\bar{x}$ pour $x \in S^{n-1}$, il vient $u^{-1} = \varepsilon^m\bar{u}$ et, par suite,

$$u^{-1} = \bar{u} \quad \text{si} \quad u \in \mathrm{Spin}_\varepsilon(n) \quad (\text{ou } \varepsilon = +1).$$

Si $u = \sum_{i=1}^{n} u_ie_i \in \mathbb{R}^n$ et $x = \sum_{i=1}^{n} x_ie_i \in \mathbb{R}^n$, alors

$$ux = -\sum_{i=1}^{n} u_ix_i + \sum_{i \neq j} u_ix_je_ie_j,$$

et, par suite,

$$ux\bar{u} = \sum_{(ijk)} u_ix_ju_ke_ie_je_k + \ldots,$$

où les points de suspension figurent des termes linéaires en $e_1, \ldots, e_n$ (c'est-à-dire appartenant à $\mathbb{R}^n$) et le symbole placé sous le signe somme que la sommation est étendue à tous les triples (ijk) de nombres $1, \ldots, n$ deux à deux distincts. Or, le produit $e_ie_je_k$ change de signe par une permutation de deux quelconques de ses facteurs et les coefficients $u_ix_ju_k$ sont symétriques par rapport à i et k. Donc, cette somme est nulle et, par suite, $ux\bar{u} \in \mathbb{R}^n$. Comme $(uv)x(\overline{uv}) = u(vx\bar{v})\bar{u}$, on en déduit que $ux\bar{u} \in \mathbb{R}^n$ pour tout élément $u \in \mathrm{Cl}(n)$

représentable par un produit d'éléments de $\mathbb{R}^n$ et donc, en particulier, pour tout élément $u \in \text{pin}_\varepsilon(n)$. Ceci montre qu'en posant

$$\varphi(u)x = ux\bar{u}, \quad u \in \text{pin}_\varepsilon(n), \quad x \in \mathbb{R}^n,$$

on obtient une application (visiblement linéaire)

$$\varphi(u)\colon \mathbb{R}^n \to \mathbb{R}^n.$$

On voit par ailleurs que $\varphi(uv) = \varphi(u)\varphi(v)$, c'est-à-dire que l'application $\varphi\colon u \mapsto \varphi(u)$ est un homomorphisme du groupe $\text{pin}_\varepsilon(n)$ dans le groupe des opérateurs linéaires inversibles de $\mathbb{R}^n$ dans $\mathbb{R}^n$, groupe que, compte tenu du fait que nous avons fixé une base $e_1, \ldots, e_n$ dans $\mathbb{R}^n$, nous identifierons au groupe $GL(n)$ des matrices inversibles.

Par ailleurs, comme pour tout élément $x \in \mathbb{R}^n$ on a l'égalité

$$|\varphi(u)x|^2 = \varepsilon(\varphi(u)x)^2 = \varepsilon ux\bar{u}ux\bar{u} =$$
$$= \varepsilon uxu^{-1}uxu^{-1} = \varepsilon uxxu^{-1} = |x|^2uu^{-1} = |x|^2,$$

l'opérateur $\varphi(u)$ est orthogonal pour tout élément $u \in \text{pin}_\varepsilon(u)$, de sorte que φ est bien un homomorphisme

$$\varphi\colon \text{pin}_\varepsilon(n) \to O(n).$$

Proposition 4. *L'application*

$$\varphi\colon \text{pin}_\varepsilon(n) \to O(n)$$

pour $\varepsilon = -1$ *ou* n *pair est un épimorphisme sur le groupe* $O(n)$, *et pour* $\varepsilon = +1$ *et* n *impair, un épimorphisme sur le groupe* $SO(n)$.

Pour tous ε *et* n, *l'homomorphisme* φ *envoie le groupe* $\text{Spin}_\varepsilon(n)$ *dans le groupe* $SO(n)$ *et, de plus, l'homomorphisme induit*

$$\varphi_0\colon \text{Spin}_\varepsilon(n) \to SO(n)$$

est un épimorphisme.

Le noyau de l'épimorphisme φ_0 *est le groupe du second ordre* $\{1, -1\}$.

Démonstration. Si $u \in S^{n-1}$ et $x \in \mathbb{R}^n$, alors d'après la formule (2) on a

$$\varphi(u)x = ux\bar{u} = uxu = (2\varepsilon(x, u) - xu)u =$$
$$= -\varepsilon(x - 2(x, u)u)$$

et, par suite,

$$\varphi(u) = -\varepsilon u^\perp,$$

où $u^\perp\colon x \mapsto x - 2(x, u)u$ est la symétrie dans l'hyperplan orthogonal au vecteur u. Donc, l'image de l'homomorphisme φ pour $\varepsilon = -1$ contient toute symétrie de l'espace $\mathbb{R}^n$, et pour $\varepsilon = 1$, le composé d'une symétrie quelconque et de l'opérateur $x \mapsto -x$ de détermi-

nant $(-1)^n$. Ceci prouve les trois premières assertions de la proposition 4, puisqu'on sait (cf. I, 27) que dans les espaces à une et à deux dimensions, tout opérateur orthogonal est une symétrie ou une composée de symétries, et dans un espace à n dimensions, la somme directe d'opérateurs orthogonaux dans des sous-espaces à une et à deux dimensions (cf. II, 21) et, par suite, est aussi une symétrie ou une composée de symétries. Ceci étant, un opérateur appartient à $SO(n)$ si et seulement s'il est le composé d'un nombre pair de symétries.

Si $u \in \operatorname{Ker} \varphi_0$, alors $ue_i = e_iu$ pour tout $i = 1, \ldots, n$ (car, comme signalé plus haut, $u^{-1} = \bar{u}$ pour tout élément $u \in \mathrm{Spin}_\varepsilon(n)$) et, par suite, u appartient au centre de l'algèbre $Cl_\varepsilon(n)$, c'est-à-dire, étant pair, est un nombre de $\mathbb{R}$. Donc, $\varphi(u)x = u^2x$, c'est-à-dire que $\varphi(u) = u^2E$ et, par suite, $u = \pm 1$ puisque l'opérateur $\varphi(u)$ est orthogonal. Réciproquement, il est clair que $\pm 1 \in \operatorname{Ker} \varphi_0$. □

De la proposition 4, il s'ensuit que si le groupe de Lie $\mathrm{Spin}_\varepsilon(n)$ est connexe pour $n > 1$, l'épimorphisme φ_0 est un revêtement de groupe. Dans le cas contraire, le groupe $\mathrm{Spin}_\varepsilon(n)$ doit être le produit direct $SO(n) \times \mathbb{Z}_2$ (et l'application φ_0, la projection $SO(n) \times \mathbb{Z}_2 \to SO(n)$), donc ses points 1 et -1 seront situés dans des composantes différentes. Or il est évident qu'en posant

$$u(t) = \varepsilon\left(\cos\frac{\pi}{2}t\cdot e_1 + \sin\frac{\pi}{2}t\cdot e_2\right)\left(\cos\frac{\pi}{2}t\cdot e_1 - \sin\frac{\pi}{2}t\cdot e_2\right) = \\ = \cos\pi t - \varepsilon\sin\pi t\cdot e_1e_2, \qquad 0 \leqslant t \leqslant 1,$$

on obtient dans $\mathrm{Spin}_\varepsilon(n)$ un chemin $t \mapsto u(t)$ reliant le point 1 au point -1. C'est donc le premier cas qui aura lieu. Par conséquent, *le groupe* $\mathrm{Spin}_\varepsilon(n)$ *est connexe pour* $n > 1$ *et l'épimorphisme* φ_0 *est un revêtement de groupe.*

L'image réciproque de tout point par le revêtement φ_0 est composée de deux éléments. De tels revêtements sont appelés *revêtements à deux feuillets.*

L'existence d'un revêtement non trivial du groupe $SO(n)$ exprime que *ce groupe n'est pas simplement connexe.*

Pour trouver son groupe fondamental, on se servira encore de la proposition 8 de la leçon précédente. Comme déjà signalé dans la leçon 1, la variété quotient $SO(n)/SO(n-1)$ s'identifie canoniquement à la sphère S^{n-1} et, par suite, est simplement connexe pour $n \geqslant 3$. Donc, d'après la proposition 8 de la leçon 12, le groupe fondamental $\pi_1 SO(n)$ du groupe $SO(n)$ pour tout $n \geqslant 3$ est le groupe quotient du groupe $\pi_1 SO(3)$. Il suffit donc de calculer le groupe $\pi_1 SO(3)$.

Soient $\mathbb{H}'$ l'espace vectoriel des quaternions « imaginaires purs » η (c'est-à-dire tels que $\bar{\eta} = -\eta$) et S^3, le groupe des quaternions « unitaires » ξ (c'est-à-dire tels que $\bar{\xi} = \xi^{-1}$). Si $\xi \in S^3$ et $\eta \in \mathbb{H}'$,

alors $\overline{\xi\eta\xi^{-1}} = \bar{\xi}^{-1}\bar{\eta}\,\bar{\xi} = -\xi\eta\xi^{-1}$ et, par suite, $\xi\eta\xi^{-1} \in H'$. Donc, pour tout quaternion $\xi \in S^3$, la formule $\varphi(\xi): \eta \mapsto \xi\eta\xi^{-1}$, $\eta \in H'$, définit un opérateur (visiblement linéaire) $\varphi(\xi): H' \to H'$. Ceci étant, l'opérateur $\varphi(\xi)$ est orthogonal, puisque $|\xi\eta\xi^{-1}| = |\eta|$. En vertu de l'identification $H' = \mathbb{R}^3$, l'application $\varphi: \xi \mapsto \varphi(\xi)$ sera donc une application (de toute évidence un homomorphisme) de S^3 dans O(3). Bien plus, le groupe S^3 étant connexe, l'homomorphisme φ est bien une application dans le groupe SO(3).

Proposition 5. *L'homomorphisme*

$$\psi: S^3 \to SO(3)$$

est un épimorphisme. Son noyau est le groupe du deuxième ordre $\{1, -1\}$.

Pour prouver cette proposition on se servira du lemme général suivant:

Lemme 2. *L'homomorphisme* $\Phi: G \to H$ *de groupes de Lie connexes est un revêtement de groupe (et, par suite, un épimorphisme) si son noyau* $K = \operatorname{Ker}\Phi$ *est discret et* $\dim G = \dim H$.

D é m o n s t r a t i o n. Ayant un noyau discret, l'homomorphisme Φ, considéré comme une application sur $\Phi(G) \approx G/K$, est un revêtement de groupe. Donc, seule l'égalité $\Phi(G) = H$ est à démontrer. Mais comme $\dim \Phi(G) = \dim G = \dim H$, l'unité $e \in \Phi(G)$ est un point intérieur du sous-groupe $\Phi(G)$, c'est-à-dire que dans H il existe un voisinage U de l'unité tel que $U \subset \Phi(G)$. Ceci prouve le lemme puisque le groupe H étant connexe, il est engendré par le voisinage U. □

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 4. Comme $\dim S^3 = \dim SO(3) = 3$, il suffit en vertu du lemme 2 de prouver seulement l'assertion relative au noyau. Or, si $\xi \in \operatorname{Ker}\psi$, pour les unités quaternioniques i, j, k, on aura les égalités $\xi i = i\xi$, $\xi j = j\xi$, $\xi k = k\xi$, ce qui n'est possible de toute évidence que si $\xi \in \mathbb{R}$, c'est-à-dire pour $\xi = \pm 1$. □

La proposition 4 dit que l'application $\psi: S^3 \to SO(3)$ est un revêtement de groupe à deux feuillets. La sphère S^3 étant simplement connexe, ce revêtement est universel. Donc, $\pi_1 SO(3) = \mathbb{Z}_2$ et, par suite, d'après la remarque faite plus haut, le groupe $\pi_1 SO(n)$ est un groupe quotient du groupe $\mathbb{Z}_2$ pour tout $n \geqslant 3$. Or nous savons que ce groupe n'est pas trivial. Donc, *le groupe* $\pi_1 SO(n)$ *est le groupe* $\mathbb{Z}_2$ *du deuxième ordre pour* $n \geqslant 3$.

Pour $n = 2$, le groupe de Lie SO(2) est un cercle S^1, donc *le groupe* $\pi_1 SO(2)$ *est isomorphe au groupe* $\mathbb{Z}$.

Pour $n = 1$ le groupe SO(1) est un groupe composé uniquement de l'unité.

Par ailleurs, nous constatons maintenant que *le groupe* $\mathrm{Spin}_\varepsilon(n)$ *est un groupe simplement connexe pour* $n \geqslant 3$, *et le revêtement* φ_0: $\mathrm{Spin}_\varepsilon(n) \to \mathrm{SO}(n)$, *un revêtement universel.*

Ce revêtement universel étant unique, on déduit, en particulier, que *le groupe* $\mathrm{Spin}_+(n)$ *est isomorphe au groupe* $\mathrm{Spin}(n)$.

Ainsi, les algèbres $\mathrm{Cl}_+(n)$ et $\mathrm{Cl}(n)$ ne sont pas isomorphes, mais leurs sous-groupes $\mathrm{Spin}_+(n)$ et $\mathrm{Spin}(n)$ le sont.

Remarque 2. On voudrait bien avoir l'isomorphisme des groupes $\mathrm{Spin}_+(n)$ et $\mathrm{Spin}(n)$ sous une forme plus explicite. De même qu'on voudrait tout aussi bien appréhender le sens algébrique de cet isomorphisme. Nos vœux seront exaucés si l'on construit un isomorphisme des algèbres $\mathrm{Cl}_+^0(n)$ et $\mathrm{Cl}^0(n)$ appliquant le groupe $\mathrm{Spin}_+(n)$ sur le groupe $\mathrm{Spin}(n)$. Cet isomorphisme sera un isomorphisme linéaire $\rho: \mathrm{Cl}_+^0(n) \to \mathrm{Cl}^0(n)$ envoyant élément de base e_I de l'algèbre $\mathrm{Cl}_+^0(n)$ dans élément de base $\bar{e}_I$ de l'algèbre $\mathrm{Cl}^0(n)$, c'est-à-dire dans $(-1)^p e_I$, où $2p = |I|$. En effet, il suffit de toute évidence de montrer que ρ est un homomorphisme d'algèbres, c'est-à-dire que $\rho(e_I e_J) = \rho(e_I)\rho(e_J)$ pour tous éléments de base e_I, e_J de l'algèbre $\mathrm{Cl}_+^0(n)$. Soient $|I| = 2p$, $|J| = 2q$ et $|I \triangle J| = 2r$. Soient par ailleurs τ_+ le nombre de couples $(i, j) \in I \times J$ tels que $i > j$ et τ, celui des couples tels que $i \geqslant j$. D'après les formules (12) et (13) on a $e_I e_J = (-1)^{\tau_+} e_{I \triangle J}$ dans $\mathrm{Cl}_+^0(n)$ et $e_I e_J = (-1)^{\tau} e_{I \triangle J}$ dans $\mathrm{Cl}^0(n)$. Donc, $\rho(e_I e_J) = (-1)^{\tau_+ + s} e_{I \triangle J}$, et $\rho(e_I)\rho(e_J) = (-1)^{\tau + p + q} e_{I \triangle J}$. Ce que nous voulions, puisque de toute évidence $\tau - \tau_+ = |I \cap J| = p + q - s$. □

Remarque 3. Le fait que l'application ρ est un homomorphisme d'algèbres peut être prouvé sans calculs si l'on remarque qu'elle est la restriction d'un isomorphisme des algèbres complexifiées $\mathrm{Cl}_+(n) \otimes \mathbb{C}$ et $\mathrm{Cl}(n) \otimes \mathbb{C}$, induit par les correspondances $e_1 \mapsto ie_1, \ldots, e_n \mapsto ie_n$, où $i = \sqrt{-1}$.

Dans la suite nous n'étudierons en principe que le groupe $\mathrm{Spin}(n)$.

La description des groupes $\mathrm{Spin}(n)$ pour n petit ne pose aucun problème.

Il est clair que le groupe $\mathrm{Spin}(1)$ est, tout comme le groupe $\mathrm{SO}(1)$, composé seulement de l'unité.

L'algèbre $\mathrm{Cl}^0(2)$ est à deux dimensions (une base est constituée des éléments 1 et e_1e_2), et le groupe $\mathrm{Spin}(2)$ est un cercle dans cet espace vectoriel à deux dimensions. Donc,

$$\mathrm{Spin}(2) \approx \mathrm{SO}(2) \approx \mathbb{S}^1 \approx \mathrm{U}(1).$$

S'agissant du groupe $\mathrm{Spin}(3)$, il est isomorphe au groupe $\mathbb{S}^3$ en vertu de l'unicité du revêtement universel:

$$\mathrm{Spin}(3) \approx \mathbb{S}^3 \approx \mathrm{Sp}(1).$$

Il est intéressant de constater que *le groupe* Spin(3) $\approx \mathcal{S}^3$ *est également isomorphe au groupe* SU(2). En effet, un calcul immédiat montre que toute matrice de SU(2) est de la forme

$$(14) \qquad \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \text{ où } |a|^2 + |b|^2 = 1,$$

et que l'application $SU(2) \to \mathcal{S}^3$ qui associe à la matrice (14) le quaternion $\xi = a + bj \in \mathcal{S}^3$ est un isomorphisme. □

On en déduit, en particulier, que *le groupe* SU(2) *est un revêtement à deux feuillets du groupe des rotations* SO(3).

Pour obtenir l'expression explicite du revêtement $SU(2) \to$ $\to SO(3)$ associons à chaque matrice (14) une transformation homographique

$$(15) \qquad z \to \frac{az + b}{-\bar{b}z + \bar{a}}$$

du plan complexe achevé $\mathbf{C}^+$. Si le plan $\mathbf{C}^+$ est identifié à une sphère $\mathcal{S}^2$ au moyen d'une projection stéréographique, les transformations (15) sont envoyées (cf. I, 28) dans les rotations de la sphère, c'est-à-dire dans des transformations de SO(3). Ceci nous donne le revêtement $SU(2) \to SO(3)$, puisque les matrices qui ne diffèrent que par leur signe engendrent la même rotation (15).

Par ailleurs, il est immédiat de voir que *le groupe* Spin(4) *est isomorphe au produit direct* $\mathcal{S}^3 \times \mathcal{S}^3$. Le moyen le plus simple pour le prouver est de remarquer, *primo*, que pour tout couple de quaternions $\xi, \eta \in \mathcal{S}^3$, la formule

$$\zeta \to \xi\zeta\bar{\eta}, \quad \zeta \in \mathbf{H},$$

définit une isométrie (puisque $|\xi\zeta\bar{\eta}| = |\xi|\cdot|\zeta|\cdot|\bar{\eta}| = |\zeta|$) de l'algèbre $\mathbf{H}$ dans elle-même, c'est-à-dire un élément du groupe SO(4), *secundo*, que l'application obtenue $\mathcal{S}^3 \times \mathcal{S}^3 \to SO(4)$ est un homomorphisme (car $\xi_1(\xi_2\zeta\bar{\eta}_2)\bar{\eta}_1 = (\xi_1\xi_2)\,\zeta\overline{(\eta_1\eta_2)}$), et, *tertio*, que le noyau de cet homomorphisme est composé seulement des deux éléments (1, 1) et (—1, —1) (si $\xi\zeta\bar{\eta} = \zeta$ pour tous les ζ, alors, en particulier, $\xi\bar{\eta} = 1$ et, par suite, $\xi = \eta$; donc $\xi\zeta = \zeta\xi$, d'où, comme on le sait, il s'ensuit que $\xi = \pm 1$). Cela signifie (cf. lemme 1) que le groupe $\mathcal{S}^3 \times \mathcal{S}^3$ est un revêtement à deux feuillets du groupe SO(4). Donc, ce revêtement est universel et, par suite, le groupe $\mathcal{S}^3 \times \mathcal{S}^3$ est isomorphe au groupe Spin(4). □

Des résultats analogues caractérisent les groupes Spin(5) et Spin (6). Pour les obtenir, nous commençons par le groupe $SL(4; \mathbf{C})$ des opérateurs linéaires unimodulaires de l'espace complexe à quatre dimensions $\mathbf{C}^4$. Chaque opérateur $A \in SL(4;\mathbf{C})$ induit de manière évi-

dente un opérateur linéaire $\hat{A}$. $\Lambda^2(\mathbb{C}^4) \to \Lambda^2(\mathbb{C}^4)$ dans l'espace vectoriel $\Lambda^2(\mathbb{C}^4)$ des fonctionnelles antisymétriques bilinéaires de $\mathbb{C}^4$ de type (0, 2). Cet opérateur agit sur les bivecteurs à l'aide de la formule

$$\hat{A}(x \wedge y) = Ax \wedge Ay, \quad x, y \in \mathbb{C}^4.$$

Comme dim $\Lambda^2(\mathbb{C}^4) = 6$, en choisissant dans l'espace $\Lambda^2(\mathbb{C}^4)$ une base composée des bivecteurs $e_i \wedge e_j$, $i < j$, où e_1, e_2, e_3, e_4 sont les vecteurs de la base canonique de l'espace $\mathbb{C}^4$, on peut considérer que l'opérateur $\hat{A}$ est un élément du groupe GL(6; $\mathbb{C}$). Bien plus, en étudiant sur l'espace $\Lambda^2(\mathbb{C}^4)$ la fonctionnelle quadratique Q qui à tout vecteur $p^{ij}e_i \wedge e_j$ de cet espace associe un nombre $p^{12}p^{34} + p^{23}p^{14} + p^{13}p^{42}$, on trouve par un calcul peu compliqué que tout opérateur $\hat{A}$ préserve Q. (Du reste ce fait peut être établi géométriquement sans le moindre calcul. En effet, l'opérateur $\hat{A}$ envoyant bivecteur dans bivecteur respecte l'égalité $Q = 0$ qui est équivalente à la relation de Plücker qui caractérise les bivecteurs parmi les fonctionnelles de $\Lambda^2(\mathbb{C}^4)$; cf. II, 10. Cela signifie que l'opérateur $\hat{A}$ applique l'hypersurface du second ordre $Q = 0$ dans elle-même. Alors, d'après le théorème d'unicité, aux coefficients de proportionnalité près, des équations des hypersurfaces du second ordre, l'opérateur $\hat{A}$ doit transformer la fonctionnelle Q en une fonctionnelle proportionnelle $\lambda_A Q$. Il suffit donc de montrer que $\lambda_A = 1$ pour tout opérateur $A \in$ SL(4; $\mathbb{C}$). Or la correspondance $A \mapsto \lambda_A$ est visiblement un homomorphisme du groupe SL(4; $\mathbb{C}$) dans le groupe multiplicatif $\mathbb{C}^*$ des nombres complexes non nuls, et on démontre sans peine qu'un tel homomorphisme est trivial.)

Si de la base $e_i \wedge e_j$, $i < j$, on passe à la base

$$(16) \qquad \begin{aligned} f_1 &= e_1 \wedge e_2 + e_3 \wedge e_4, & f_2 &= ie_1 \wedge e_2 - ie_3 \wedge e_4, \\ f_3 &= e_2 \wedge e_3 + e_1 \wedge e_4, & f_4 &= ie_2 \wedge e_3 - ie_1 \wedge e_4, \\ f_5 &= e_1 \wedge e_3 + e_4 \wedge e_2, & f_6 &= ie_1 \wedge e_3 - ie_4 \wedge e_2, \end{aligned}$$

dans laquelle la fonctionnelle Q se représente par une somme de carrés, alors chaque opérateur $\hat{A}$ sera défini dans cette base par une matrice orthogonale. En désignant cette matrice par $\chi(A)$, on obtient une application (qui est de toute évidence un homomorphisme)

$$\chi\colon \mathrm{SL}(4;\ \mathbb{C}) \to \mathrm{O}(6;\ \mathbb{C}).$$

Le noyau de l'homomorphisme χ est composé de matrices A dont les colonnes a_1, a_2, a_3, a_4 satisfont la relation $a_i \wedge a_j = e_i \wedge e_j$ quels que soient i et j et, par suite, sont telles que pour tout couple (i, j) les vecteurs a_i et a_j s'expriment linéairement en fonction des vecteurs e_i et e_j, ce qui évidemment est possible seulement si $a_i = \lambda_i e_i$ pour

tout i. Par ailleurs, pour tous i et j, on doit avoir l'égalité $\lambda_i\lambda_j = 1$, ce qui n'est possible que si ou bien $\lambda_i = 1$ ou bien $\lambda_i = -1$ pour tout i. Ceci prouve que *le noyau de l'homomorphisme* χ *est le groupe du second ordre* $\{E, -E\}$.

Considérons maintenant le sous-groupe de $SL(4; C)$ composé des matrices A telles que la matrice $\chi(A)$ possède des coefficients réels, c'est-à-dire appartient au groupe $O(6) = O(6; R)$. Il est clair qu'une matrice A appartient à ce sous-groupe si et seulement si l'opérateur $\hat{A}$ commute à la transformation semi-linéaire S: $\Lambda^2(C^4) \to \Lambda^2(C^4)$ qui change les coordonnées de chaque élément de l'espace vectoriel $\Lambda^2(C^4)$ dans la base (16) en leurs conjuguées complexes. Si une fonctionnelle de $\Lambda^2(C^4)$ admet les coordonnées $z_1, \ldots, z_6$ dans la base (16), elle admettra visiblement les coordonnées suivantes

$$p_{12} = z_1 + iz_2, \quad p_{34} = z_1 - iz_2,$$
$$p_{23} = z_3 + iz_4, \quad p_{14} = z_3 - iz_4,$$
$$p_{13} = z_5 + iz_6, \quad p_{24} = -z_5 + iz_6,$$

dans la base composée des bivecteurs $e_i \wedge e_j$, $i < j$, donc sa transformée par S sera la fonctionnelle de coordonnées

$$\overline{p}_{34}, \quad \overline{p}_{12}, \quad \overline{p}_{14}, \quad \overline{p}_{23}, \quad -\overline{p}_{24}, \quad -\overline{p}_{13}.$$

Cela signifie que $S = T \circ \hat{\sigma}$, où T est un opérateur linéaire $\Lambda^2(C^4) \to \Lambda^2(C^4)$ agissant sur les bivecteurs de base à l'aide des formules

$$\begin{aligned} &T(e_1 \wedge e_2) = e_3 \wedge e_4, && T(e_3 \wedge e_4) = e_1 \wedge e_2, \\ (17) \qquad &T(e_2 \wedge e_3) = e_1 \wedge e_4, && T(e_1 \wedge e_4) = e_2 \wedge e_3, \\ &T(e_1 \wedge e_3) = -e_2 \wedge e_4, && T(e_2 \wedge e_4) = -e_1 \wedge e_3, \end{aligned}$$

et $\sigma: C^4 \to C^4$ un isomorphisme semi-linéaire remplaçant les composantes de chaque vecteur par leurs conjuguées complexes. Donc, $\chi(A) \in O(6)$ si et seulement si

$$\hat{A} \circ T \circ \hat{\sigma} = T \circ \hat{\sigma} \circ \hat{A}.$$

On montrera plus bas que tout opérateur $A \in SL(4; C)$ est justiciable de la relation

$$(18) \qquad \hat{A} \circ T = T \circ \hat{A}^c,$$

où A^c est un opérateur sur C^4 dont la matrice se déduit de celle de l'opérateur A par une transposition et une inversion. Il s'ensuit de là que $\chi(A) \in O(6)$ si et seulement si $\hat{A}^c \circ \hat{\sigma} = \hat{\sigma} \circ \hat{A}$, c'est-à-dire si et seulement si $\widehat{A^c \circ \sigma} = \widehat{\sigma \circ A}$. En particulier, $\chi(A) \in O(6)$ si $A^c \circ \sigma = \sigma \circ A$, c'est-à-dire si $A^* = A^{-1}$, où A^* est l'opérateur

adjoint (pour la multiplication scalaire standard de $\mathbb{C}^4$) (la matrice de A^* est la transposée conjuguée de celle de A). Comme l'égalité $A^* = A^{-1}$ caractérise les opérateurs unitaires appartenant au sous-groupe SU(4) du groupe SL(4; $\mathbb{C}$), ceci prouve que l'homomorphisme χ envoie le groupe SU(4) dans le groupe O(6) et, puisque le groupe SU(4) est connexe, dans le groupe SO(6). Il induit donc l'homomorphisme

$$\chi_0 : SU(4) \to SO(6).$$

Cet homomorphisme est un revêtement à deux feuillets, puisque son noyau est confondu de toute évidence avec celui de χ, et donc est un groupe du second ordre, et dim SU(4) = dim SO(6) = 15. Ceci prouve que *le groupe* SU(4) *est un revêtement à deux feuillets du groupe* SO(6) *et, par suite, est isomorphe au groupe* Spin(6).

Le groupe SU(4) contient le sous-groupe Sp(2) dont les éléments A sont caractérisés par le fait qu'ils laissent invariante la fonctionnelle bilinéaire antisymétrique de matrice

$$J = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

ou, ce qui est équivalent, par le fait que $A^c J = JA$, où J est un opérateur linéaire sur $\mathbb{C}^4$ de matrice J (comparer avec la formule (5) de la leçon 1). Donc, si $A \in Sp(2)$, alors $\hat{A}^c \circ \hat{J} = \hat{J} \circ \hat{A}$ et, par suite, (cf. formule (18)), $\hat{A} \circ T \circ \hat{J} = T \circ \hat{J} \circ \hat{A}$. Mais un calcul évident montre que $T \circ \hat{J} = -I$, où I est un opérateur linéaire $\Lambda^2(\mathbb{C}^4) \to \Lambda^2(\mathbb{C}^4)$ laissant invariants tous les vecteurs (16), hormis le vecteur f_3 qui est envoyé dans $-f_3$. Donc, $\hat{A} \circ I = I \circ \hat{A}$, ce qui équivaut à l'égalité $\hat{A}f_3 = \pm f_3$. Ceci signifie que $\chi(A)$ appartient à un sous-groupe de SO(6) qui est isomorphe au groupe O(5) et même — puisque Sp(2) est connexe — à la composante de l'unité SO(5) de ce sous-groupe. Comme dim Sp(2) = dim SO(5) = 10, il vient que *le groupe* Sp(2) *est un revêtement à deux feuillets du groupe* SO(5), *et, par conséquent, est isomorphe au groupe* Spin(5).

En rassemblant tous les faits établis, on obtient la proposition suivante:

Proposition 6. *Les groupes*

$$S^1, \quad S^3 = Sp(1) = SU(2), \quad S^3 \times S^3, \quad Sp(2), \quad SU(4)$$

sont des revêtements à deux feuillets des groupes

$$SO(2), \quad SO(3), \quad SO(4), \quad SO(5), \quad SO(6)$$

et, par conséquent, sont isomorphes aux groupes

$$\mathrm{Spin}(2), \quad \mathrm{Spin}(3) \quad \mathrm{Spin}(4), \quad \mathrm{Spin}(5), \quad \mathrm{Spin}(6)$$

respectivement. □

Nous avons ainsi réussi à représenter les groupes $\mathrm{Spin}(n)$ pour $n \leqslant 6$ sous forme de groupes matriciels. On obtient des résultats analogues pour $n > 6$ à ceci près que les groupes $\mathrm{Spin}(n)$, $n > 6$, s'identifient seulement à des sous-groupes orthogonaux correspondants.

Il est évident que pour prouver cette assertion, il suffit d'obtenir les représentations matricielles des algèbres complètes de Clifford $\mathrm{Cl}_\varepsilon(n)$. Dans ces représentations, nous négligerons la $\mathbb{Z}_2$-graduation et donc, en particulier, nous admettrons que tous les produits tensoriels sont ordinaires (et pas gauches).

Nous savons déjà que $\mathrm{Cl}(1) \approx \mathbb{C}$ et $\mathrm{Cl}_+(1) \approx \mathbb{D}$, où $\mathbb{D}$ est l'algèbre des nombres doubles $a + be$, $e^2 = 1$. Du reste, il est immédiat de voir que la correspondance $a + be \to (a + b, a - b)$ est un isomorphisme $\mathbb{D} \approx \mathbb{R} \oplus \mathbb{R}$. Donc,

$$\mathrm{Cl}(1) \approx \mathbb{C}, \quad \mathrm{Cl}_+(1) \approx \mathbb{R} \oplus \mathbb{R}.$$

Une vérification automatique montre par ailleurs que les correspondances $e_1 \mapsto i$, $e_2 \mapsto j$, $e_1e_2 \mapsto k$ définissent un isomorphisme entre l'algèbre $\mathrm{Cl}(2)$ et l'algèbre des quaternions $\mathbb{H}$, et les correspondances

$$e_1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e_2 \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad e_1e_2 \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

un isomorphisme entre $\mathrm{Cl}_+(2)$ et $\mathbb{R}(2)$. Donc,

$$\mathrm{Cl}(2) \approx \mathbb{H}, \quad \mathrm{Cl}_+(2) \approx \mathbb{R}(2).$$

Proposition 7. *Pour tout $n \geqslant 0$, on a l'isomorphisme*

$$\mathrm{Cl}_\varepsilon(n+2) \approx \mathrm{Cl}_{-\varepsilon}(n) \otimes \mathrm{Cl}_\varepsilon(2),$$

c'est-à-dire les deux isomorphismes

$$\mathrm{Cl}(n+2) \approx \mathrm{Cl}_+(n) \otimes \mathbb{H},$$
$$\mathrm{Cl}_+(n+2) \approx \mathrm{Cl}(n) \otimes \mathbb{R}(2).$$

Démonstration. Considérons l'application linéaire

$$\alpha\colon \mathbb{R}^{n+2} \to \mathrm{Cl}_{-\varepsilon}(n) \otimes \mathrm{Cl}_\varepsilon(2),$$

pour laquelle

$$\alpha(e_i) = e_i \otimes e_1e_2 \text{ pour } i = 1, \ldots, n,$$
$$\alpha(e_{n+1}) = 1 \otimes e_1, \quad \alpha(e_{n+2}) = 1 \otimes e_2$$

(les générateurs des trois algèbres $\mathrm{Cl}_\varepsilon(n+2)$, $\mathrm{Cl}_{-\varepsilon}(n)$ et $\mathrm{Cl}_\varepsilon(2)$ sont désignés par les mêmes symboles e_i). Une vérification immédiate montre que $\alpha(x)^2 = \varepsilon|x|^2$ pour tout $x \in \mathbb{R}^{n+2}$. Donc, l'application α se prolonge en un homomorphisme $\alpha^\#: \mathrm{Cl}_\varepsilon(n+2) \to \mathrm{Cl}_{-\varepsilon}(n) \otimes \mathrm{Cl}_\varepsilon(2)$ de l'algèbre $\mathrm{Cl}_\varepsilon(n+2)$ dans l'algèbre $\mathrm{Cl}_{-\varepsilon}(n) \otimes \mathrm{Cl}_\varepsilon(2)$. Un raisonnement analogue nous conduit aux homomorphismes $\beta^\#: \mathrm{Cl}_{-\varepsilon}(n) \to \mathrm{Cl}_\varepsilon(n+2)$ et $\gamma^\#: \mathrm{Cl}_\varepsilon(2) \to \mathrm{Cl}_\varepsilon(n+2)$ pour lesquels

$$\beta^\#(e_i) = -e_i e_{n+1} e_{n+2}, \quad i = 1, \ldots, n,$$

et

$$\gamma^\#(e_1) = e_{n+1}, \quad \gamma^\#(e_2) = e_{n+2}.$$

Comme de toute évidence $\beta^\#(e_i)\,\gamma^\#(e_j) = \gamma^\#(e_j)\,\beta^\#(e_i)$ pour tous $i = 1, \ldots, n$ et $j = 1, 2$, les homomorphismes $\beta^\#$ et $\gamma^\#$ commutent et, par suite, l'application

$$\beta^\# \otimes \gamma^\#: \mathrm{Cl}_{-\varepsilon}(n) \otimes \mathrm{Cl}_\varepsilon(2) \to \mathrm{Cl}_\varepsilon(n+2)$$

est un homomorphisme d'algèbres. Ceci étant, pour tout $i = 1, \ldots, n$ on aura

$$[(\beta^\# \otimes \gamma^\#) \circ \alpha^\#](e_i) = \beta^\#(e_i)\cdot\gamma^\#(e_1 e_2) = -e_i\,(e_{n+1}e_{n+2})^2 = e_i$$

(car $(e_1, e_2)^2 = -1$ pour tout ε), et pour tout $j = 1, 2$

$$[(\beta^\# \otimes \gamma^\#) \circ \alpha^\#](e_{n+j}) = 1 \otimes \gamma^\#(e_j) = 1 \otimes e_{n+j},$$

de sorte que $(\beta^\# \otimes \gamma^\#) \circ \alpha^\# = \mathrm{id}$. De façon analogue, pour tout $i = 1, \ldots, n$

$$\begin{aligned}[\alpha^\# \circ (\beta^\# \otimes \gamma^\#)]\,(e_i \otimes 1) &= \alpha^\#(\beta^\#(e_i)) = \alpha^\#(-e_i e_{n+1} e_{n+2}) = \\ &= -\alpha^\#(e_i)\,\alpha^\#(e_{n+1})\,\alpha^\#(e_{n+2}) = \\ &= -(e_i \otimes e_1 e_2)(1 \otimes e_1)(1 \otimes e_2) = \\ &= -e_i \otimes (e_1 e_2)^2 = e_i\end{aligned}$$

et pour tout $j = 1, 2$

$$\begin{aligned}[\alpha^\# \circ (\beta^\# \otimes \gamma^\#)]\,(1 \otimes e_j) &= \alpha^\#(\gamma^\#(e_j)) = \\ &= \alpha^\#(e_{n+j}) = 1 \otimes e_j,\end{aligned}$$

d'où il résulte aussitôt que $\alpha^\# \circ (\beta^\# \otimes \gamma^\#) = \mathrm{id}$. Donc, les applications $\alpha^\#$ et $\beta^\# \otimes \gamma^\#$ sont des isomorphismes réciproques l'un de l'autre. □

Cette proposition entraîne immédiatement

$$\mathrm{Cl}(3) \approx (\mathbb{R} \oplus \mathbb{R}) \otimes \mathbb{H} \approx \mathbb{H} \oplus \mathbb{H}, \quad \mathrm{Cl}_+(3) \approx \mathbb{C} \otimes \mathbb{R}(2) \approx \mathbb{C}(2),$$

$$\mathrm{Cl}(4) \approx \mathbb{R}(2) \otimes \mathbb{H} \approx \mathbb{H}(2), \quad \mathrm{Cl}_+(4) \approx \mathbb{H} \otimes \mathbb{R}(2) \approx \mathbb{H}(2).$$

Pour aller de l'avant, nous aurons besoin du lemme suivant.

Lemme 3. *On a les isomorphismes*

$$\mathbb{C} \otimes \mathbb{C} \approx \mathbb{C} \oplus \mathbb{C}, \quad \mathbb{C} \otimes \mathbb{H} \approx \mathbb{C}(2), \quad \mathbb{H} \otimes \mathbb{H} \approx \mathbb{R}(4).$$

D é m o n s t r a t i o n. Il est évident que $\mathbb{C} \otimes \mathrm{Cl}(n) \approx \approx \mathbb{C} \otimes \mathrm{Cl}_{+}(n)$. Donc,

$$\mathbb{C} \otimes \mathbb{C} \approx \mathbb{C} \otimes \mathrm{Cl}(1) \approx \mathbb{C} \otimes \mathrm{Cl}_{+}(1) \approx \mathbb{C} \otimes (\mathbb{R} \oplus \mathbb{R}) \approx \mathbb{C} \oplus \mathbb{C}.$$

De façon analogue

$$\mathbb{C} \otimes \mathbb{H} \approx \mathbb{C} \otimes \mathbb{C}(2) \approx \mathbb{C} \otimes \mathbb{C}_{+}(2) \approx \mathbb{C} \otimes \mathbb{R}(2) \approx \mathbb{C}(2).$$

A noter que ces isomorphismes peuvent être établis directement et sans peine. Ainsi, l'isomorphisme $\mathbb{C} \otimes \mathbb{C} \approx \mathbb{C} \oplus \mathbb{C}$ est défini par les correspondances $1 \otimes 1 \mapsto (1, 1)$, $i \otimes 1 \mapsto (i, i)$, $1 \otimes i \mapsto \mapsto (i, -i)$, $i \otimes i \mapsto (-1, 1)$.

On établit l'isomorphisme $\mathbb{H} \otimes \mathbb{H} \approx \mathbb{R}(4)$ en identifiant $\mathbb{H}$ à $\mathbb{R}^4$ et en associant à tout élément $\xi \otimes \eta$ de $\mathbb{H} \otimes \mathbb{H}$, $\xi, \eta \in \mathbb{H}$, un opérateur linéaire $\omega(\xi \otimes \eta): \mathbb{H} \to \mathbb{H}$ agissant à l'aide de la formule

$$\omega(\xi \otimes \eta)\zeta = \xi\zeta\bar{\eta}, \quad \zeta \in \mathbb{H}.$$

Une vérification immédiate montre que l'opérateur $\omega(\xi \otimes \eta)$ est défini de façon unique (c'est-à-dire si $\xi \otimes \eta = \xi' \otimes \eta'$, alors $\omega(\xi \otimes \eta) = \omega(\xi' \otimes \eta')$) et que l'application ω se prolonge de façon unique par linéarité en un homomorphisme ω de l'algèbre $\mathbb{H} \otimes \mathbb{H}$ dans l'algèbre des opérateurs linéaires $\mathbb{H} \to \mathbb{H}$, c'est-à-dire, compte tenu de l'identification de $\mathbb{H}$ à $\mathbb{H}$, dans l'algèbre matricielle $\mathbb{R}(4)$. Ceci étant, comme $i1\bar{i} = 1$, $ii\bar{i} = i$, $ij\bar{i} = -j$, $ik\bar{i} = -k$, il vient

$$\omega(i \otimes i) = E_{11} + E_{22} - E_{33} - E_{44}.$$

et comme $i1\bar{j} = -k$, $ii\bar{j} = j$, $ij\bar{j} = i$, $ik\bar{j} = -1$, on a

$$\omega(i \otimes j) = -E_{14} + E_{23} + E_{32} - E_{41}.$$

On calcule de façon analogue toutes les 16 matrices $\omega(\xi \otimes \eta)$, où ξ, $\eta = 1, i, j, k$ (il est évident que $\omega(1 \otimes 1)$ est la matrice unité $E = E_{11} + E_{22} + E_{33} + E_{44}$). Quand on aura fini les calculs, on s'apercevra que toute unité matricielle $E_{\alpha\beta}$, $1 \leqslant \alpha, \beta \leqslant 4$, peut être représentée par une combinaison linéaire de matrices $\omega(\xi \otimes \eta)$. Par exemple,

$$E_{11} = \frac{1}{4}\omega(1 \otimes 1 + i \otimes i + j \otimes j + k \otimes k)$$

et

$$E_{12} = \frac{1}{4}\omega(i \otimes 1 - 1 \otimes i + k \otimes j - j \otimes k).$$

L'application ω est donc un épimorphisme et, par suite, un isomorphisme à cause de l'égalité des dimensions des algèbres $\mathbb{H} \otimes \mathbb{H}$ et $\mathbb{R}(4)$. □

On désignera par $\mathcal{A}(n)$ l'algèbre des matrices carrées d'ordre n sur une algèbre $\mathcal{A}$, c'est-à-dire dont les éléments appartiennent à $\mathcal{A}$. Si $E_{\alpha\beta}$ désignent encore les unités matricielles, la correspondance $a \otimes E_{\alpha\beta} \mapsto aE_{\alpha\beta}$ se prolonge visiblement en un isomorphisme

$$\mathcal{A} \otimes \mathbb{R}(n) \approx \mathcal{A}(n).$$

Comme l'algèbre $\mathbb{R}(n)(m)$ des matrices d'orde m dont les éléments sont des matrices d'ordre n s'identifie canoniquement à l'algèbre $\mathbb{R}(mn)$ des matrices d'ordre mn, on en déduit, en particulier, que

$$\mathbb{R}(m) \otimes \mathbb{R}(n) \approx \mathbb{R}(mn)$$

pour tous $m, n \geqslant 0$. Donc,

$$\mathcal{A}(m) \otimes \mathcal{B}(n) \approx (\mathcal{A} \otimes \mathbb{R}(m)) \otimes (\mathcal{B} \otimes \mathbb{R}(n)) \approx$$
$$\approx (\mathcal{A} \otimes \mathcal{B}) \otimes (\mathbb{R}(m) \otimes \mathbb{R}(n)) \approx (\mathcal{A} \otimes \mathcal{B})(mn)$$

quels que soient les algèbres $\mathcal{A}$ et $\mathcal{B}$ et les nombres $m, n \geqslant 0$.

Par conséquent, en vertu du lemme 3

$$\mathbb{C}(m) \otimes \mathbb{H}(n) \approx \mathbb{C}(2mn), \quad \mathbb{H}(m) \otimes \mathbb{H}(n) \approx \mathbb{R}(4mn).$$

En revenant aux algèbres de Clifford, on trouve tout d'abord que

$$\mathrm{Cl}_+(2) \otimes \mathrm{Cl}(2) \approx \mathbb{H} \otimes \mathbb{R}(2) \approx \mathbb{H}(2).$$

Donc, en appliquant deux fois la proposition 7, on obtient les isomorphismes

$$\mathrm{Cl}(n+4) \approx \mathrm{Cl}(n) \otimes \mathbb{H}(2) \quad \text{et} \quad \mathrm{Cl}_+(n+4) \approx \mathrm{Cl}_+(n) \otimes \mathbb{H}(2),$$

d'où il s'ensuit en vertu de l'isomorphisme $\mathbb{H}(2) \otimes \mathbb{H}(2) \approx \approx \mathbb{R}(16)$ que

$$\mathrm{Cl}(n+8) \approx \mathrm{Cl}(n) \otimes \mathbb{R}(16) \approx \mathrm{Cl}(n)(16)$$

et

$$\mathrm{Cl}_+(n+8) \approx \mathrm{Cl}_+(n) \otimes \mathbb{R}(16) \approx \mathrm{Cl}_+(n)(16).$$

Ceci nous donne les algèbres $\mathrm{Cl}_\varepsilon(n)$ pour tous les n, puisque nous avons déjà calculé ces algèbres pour $n \leqslant 4$. Exprimons le résultat définitif sous la forme du théorème suivant:

Théorème 2. *On a les isomorphismes*

$$\mathrm{Cl}(8m-3) \approx \mathbb{C}(2^{4m-2}), \qquad \mathrm{Cl}_+(8m-3) \approx \mathbb{H}(2^{4m-3}) \oplus \mathbb{H}(2^{4m-3}),$$
$$\mathrm{Cl}(8m-2) \approx \mathbb{R}(2^{4m-1}), \qquad \mathrm{Cl}_+(8m-2) \approx \mathbb{H}(2^{4m-2}),$$
$$\mathrm{Cl}(8m-1) \approx \approx \mathbb{R}(2^{4m-1}) \oplus \mathbb{R}(2^{4m-1}), \qquad \mathrm{Cl}_+(8m-1) \approx \mathbb{C}(2^{4m-1}),$$

$$\begin{array}{ll}
\mathrm{Cl}(8m) \approx \mathbb{R}(2^{4m}), & \mathrm{Cl}_+(8m) \approx \mathbb{R}(2^{4m}),\\
\mathrm{Cl}(8m+1) \approx \mathbb{C}(2^{4m}), & \mathrm{Cl}_+(8m+1) \approx \mathbb{R}(2^{4m}) \oplus \mathbb{R}(2^{4m}),\\
\mathrm{Cl}(8m+2) \approx \mathbb{H}(2^{4m}), & \mathrm{Cl}_+(8m+2) \approx \mathbb{R}(2^{4m+1}),\\
\mathrm{Cl}(8m+3) \approx & \mathrm{Cl}_+(8m+3) \approx \mathbb{C}(2^{4m+1}),\\
\quad \approx \mathbb{H}(2^{4m}) \oplus \mathbb{H}(2^{4m}), & \\
\mathrm{Cl}(8m+4) \approx \mathbb{H}(2^{4m+1}), & \mathrm{Cl}_+(8m+4) \approx \mathbb{H}(2^{4m+1}). \quad \square
\end{array}$$

On remarquera que

$$\mathrm{Cl}(8m) \approx \mathrm{Cl}_+(8m), \quad \mathrm{Cl}(8m+4) \approx \mathrm{Cl}_+(8m+4).$$

Il n'existe pas d'autres algèbres $\mathrm{Cl}(n)$ et $\mathrm{Cl}_+(n)$ qui soient isomorphes.

A noter aussi que

$$\mathbb{R}(2^n) = \begin{cases} \mathrm{Cl}(2n) & \text{pour } n = 4m-1,\ 4m, \\ \mathrm{Cl}_+(2n) & \text{pour } n = 4m+1. \end{cases}$$

Idem pour $\mathbb{C}(2^n)$ et $\mathbb{H}(2^n)$. Il est intéressant de signaler que l'algèbre $\mathbb{R}(2^{4m+2})$ n'est isomorphe à aucune algèbre $\mathrm{Cl}_\varepsilon(n)$.

Du théorème 2 il s'ensuit que pour les groupes $\mathrm{Spin}(n)$ (et $\mathrm{pin}_\varepsilon(n)$) il existe des monomorphismes de ces groupes dans les groupes matriciels correspondants ou dans leurs sommes directes. Mais, comme tout couple (A, B) de matrices d'ordre n s'identifie à la matrice

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \tag{19}$$

d'ordre $2n$, les injections dans les sommes directes peuvent être remplacées par des injections dans les groupes de matrices d'ordre double. Donc, *les éléments des groupes* $\mathrm{Spin}(n)$ *sont des matrices d'ordre* N *sur les algèbres* $\mathbb{R}$, $\mathbb{C}$ *ou* $\mathbb{H}$. *De plus* $N = 2^{\alpha(n)}$, *où*

$$\alpha(n) = \begin{cases} 4m-2 & \text{pour } n = 8m-3, \\ 4m-1 & \text{pour } n = 8m-2, \\ 4m & \text{pour } n = 8m-1,\ 8m,\ 8m+1,\ 8m+2, \\ 4m+1 & \text{pour } n = 8m+3,\ 8m+4, \end{cases}$$

les matrices étant réelles pour $n = 8m-2,\ 8m-1,\ 8m$, *complexes pour* $n = 8m-3,\ 8m+1$ *et quaternioniques pour* $n = 8m+2,\ 8m+3,\ 8m+4$.

Du reste, puisque l'identification du couple (A, B) à la matrice (19) se solde par une certaine perte d'information, il serait plus logique d'éviter cette identification et de représenter les éléments des

groupes Spin(n) pour $n = 8m - 1$ et $n = 8m + 3$ par des couples de matrices d'ordre $2^{\alpha(n)-1}$ (resp. réelles et quaternioniques).

Définition 5. Deux représentations matricielles d'un groupe G, c'est-à-dire deux monomorphismes (ou, de façon plus générale, deux homomorphismes) d'un groupe G dans un groupe matriciel, sont dites *équivalentes* si elles diffèrent d'un automorphisme intérieur de ce groupe, c'est-à-dire si elles s'obtiennent à partir de la même représentation du groupe G comme groupe des opérateurs linéaires d'un espace vectoriel dans des bases différentes de cet espace.

Il importe de noter que *les représentations des groupes* Spin(n) *construites plus haut sont définies à l'équivalence près*, puisque les isomorphismes de la proposition 7 et du lemme 2 peuvent être construits de plusieurs manières ne présentant aucun avantage intrinsèque par rapport l'une à l'autre (on peut admettre par exemple que $\alpha(e_1) = 1 \otimes e_1$, $\alpha(e_2) = 1 \otimes e_2$ et $\alpha(e_i) = e_{i-2} \otimes e_1 e_2$ pour $2 \leqslant i \leqslant n + 2$).

Cette assertion peut être précisée.

On dira qu'une matrice (sur une algèbre unitaire $\mathcal{A}$) est *monomiale* si dans chaque colonne et dans chaque ligne elle a tous ses éléments nuls à l'exception d'un seul qui est égal à ± 1. L'opérateur défini par une matrice monomiale permute les vecteurs de base en les multipliant simultanément par ± 1. Donc, toute matrice monomiale $\mathcal{A}$ est orthogonale pour $\mathcal{A} = \mathbb{R}$, unitaire pour $\mathcal{A} = \mathbb{C}$ et symplectique pour $\mathcal{A} = \mathbb{H}$.

Etant donné que dans les isomorphismes de la proposition 7 et du lemme 2 on peut de toute évidence limiter l'arbitraire aux permutations des éléments des bases et à la multiplication de ces éléments par ± 1, on voit que *les représentations des groupes* Spin(n) *peuvent être considérées comme étant définies aux équivalences près réalisées par les matrices monomiales.*

Il est clair que quelles que soient les matrices monomiales $A \in \mathbb{R}(m)$ et $B \in \mathbb{R}(n)$, la matrice de $\mathbb{R}(mn) \approx \mathbb{R}(m) \otimes \mathbb{R}(n)$ associée au produit tensoriel $A \otimes B$ des matrices A et B (à propos, cette matrice s'appelle *produit kroneckerien* des matrices A et B) est aussi monomiale.

Par ailleurs, il est immédiat de voir que les isomorphismes $\mathbb{C} \otimes \mathbb{H} \approx \mathbb{C}(2)$ et $\mathbb{H} \otimes \mathbb{H} \approx \mathbb{R}(4)$ du lemme 2 envoient les générateurs $\xi \otimes \eta$, où $\xi = 1, i$ ou $\xi = 1, i, j, k$ et $\eta = 1, i, j, k$, dans des matrices monomiales. Ceci est valable aussi pour l'isomorphisme $\mathrm{Cl}_+(2) \approx \mathbb{R}(2)$ (relativement aux générateurs e_1 et e_2). Par une récurrence évidente, il s'ensuit de là que *les isomorphismes du théorème 2 envoient les générateurs $e_1, \ldots, e_n$ des algèbres* Cl(n) *dans des matrices monomiales $E_1, \ldots, E_n$ (ou des couples des matrices monomiales).*

En particulier, si les matrices $E_1, \ldots, E_n$ sont réelles (resp. complexes, quaternioniques), alors elles sont orthogonales (resp. unitaires, symplectiques). En d'autres termes, ces matrices satisfont l'identité $U\bar{U}^{\top} = E$, où E est une matrice unité. Mais comme $e_i^2 = -1$, alors $E_i^2 = -E$ et, par suite, $\bar{E}_i^{\top} = -E_i$. Donc $\bar{U}^{\top} = -U$ pour toute matrice U de la forme E_i, et, par conséquent, par linéarité, pour toute matrice U de la forme u^iE_i représentant un élément arbitraire $u = u^ie_i \in \mathbb{R}^n \subset \mathrm{Cl}(n)$. Par ailleurs, comme $u^2 = -|u|^2$, alors $U^2 = -|u|^2E$ et, en particulier, $U^2 = -E$ pour $u \in S^{n-1}$. Donc, $U\bar{U}^{\top} = E$, c'est-à-dire que la matrice U est orthogonale (resp. unitaire, symplectique).

En posant, pour simplifier,

$$\mathrm{O}_{\mathbb{K}}(n) = \begin{cases} \mathrm{O}(n) & \text{pour } \mathbb{K} = \mathbb{R}, \\ \mathrm{U}(n) & \text{pour } \mathbb{K} = \mathbb{C}, \\ \mathrm{Sp}(n) & \text{pour } \mathbb{K} = \mathbb{H}, \end{cases}$$

on trouve donc que *tout élément* $u \in S^{n-1} \subset \mathrm{Cl}(n)$ *se représente par une matrice* U *appartenant au groupe* $\mathrm{O}_{\mathbb{K}}(n)$, *où* $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$ *ou* $\mathbb{H}$ *selon la valeur prise par* n.

Comme les éléments de S^{n-1} engendrent le groupe $\mathrm{pin}(n)$, cette conclusion est valable pour tout élément $u \in \mathrm{pin}(n)$ et, en particulier, pour tout élément $u \in \mathrm{Spin}(n)$.

Donc, *le groupe* $\mathrm{Spin}(n)$ *est plongé dans le groupe* $\mathrm{O}_{\mathbb{K}}(2^{\alpha(n)})$, *où*

$$\mathbb{K} = \begin{cases} \mathbb{R} \ \textit{pour} \ n = 8m-2, \ 8m-1, \ 8m, \\ \mathbb{C} \ \textit{pour} \ n = 8m-3, \ 8m+1, \\ \mathbb{H} \ \textit{pour} \ n = 8m+2, \ 8m+3, \ 8m+4. \end{cases}$$

Ceci étant, lorsque $\mathbb{K} = \mathbb{R}$, on peut affirmer de plus que $\mathrm{Spin}(n) \subset \subset \mathrm{SO}(n)$, puisque $\mathrm{Spin}(n)$ est connexe.

Lorsque $\mathbb{K} \neq \mathbb{R}$, on peut certes passer aux matrices réelles (nécessairement orthogonales et unimodulaires) en identifiant $\mathbb{C}$ à $\mathbb{R}^2$ et $\mathbb{H}$ à $\mathbb{R}^4$, mais cela ne fait pas que doubler ou quadrupler la dimension, cela entraîne une importante perte d'information.

Pour $n = 8m - 1$ ou $n = 8m + 3$, on représente en fait les éléments du groupe $\mathrm{Spin}(n)$ par des couples de matrices d'ordre deux fois moindre.

A noter que nous avons dédaigné la possibilité d'obtenir d'autres représentations à l'aide de l'algèbre $\mathrm{Cl}_+(n)$ et de l'isomorphisme $\mathrm{Spin}(n) \approx \mathrm{Spin}_+(n)$. La raison est que ces représentations n'apportent rien de nouveau et conduisent à des représentations équivalentes ou à des représentations déduites d'elles par décomplexification, complexification ou construction de la somme directe.

Il existe néanmoins une autre possibilité de construction des représentations des groupes Spin(n) qui, même si elle n'apporte rien de nouveau, permet de préciser la structure des représentations déjà construites.

Cette possibilité repose sur le fait que par définition le groupe Spin(n) est contenu dans la sous-algèbre $\mathrm{Cl}^0(n)$ de l'algèbre Cl (n) composée des éléments pairs.

Proposition 8. *L'algèbre* $\mathrm{Cl}^0(n)$ *est isomorphe à l'algèbre* $\mathrm{Cl}(n-1)$ *pour tout* $n \geqslant 1$.

D é m o n s t r a t i o n. Définissons une application linéaire

$$\omega : \mathrm{Cl}(n-1) \to \mathrm{Cl}^0(n)$$

en posant

$$\omega(e_I) = \begin{cases} e_I & \text{si } |I| \text{ est pair,} \\ e_I e_n & \text{si } |I| \text{ est impair,} \end{cases}$$

pour tout élément de base e_I de l'algèbre $\mathrm{Cl}(n-1)$, où I est un sous-ensemble quelconque de l'ensemble $[n-1] = \{1, \ldots, n-1\}$. En d'autres termes, $\omega(e_I) = e_{I^+}$, où

$$I^+ = \begin{cases} I & \text{si } |I| \text{ est pair,} \\ I \cup \{n\} & \text{si } |I| \text{ est impair.} \end{cases}$$

En désignant comme plus haut par $\tau(I, J)$ le nombre de couples $(i, j) \in I \times J$ tels que $i \geqslant j$, on trouve aussitôt que

$$\tau(I^+, J^+) \equiv \tau(I, J) \bmod 2.$$

Il est aisé de voir par ailleurs que $I^+ \triangle J^+ = (I \triangle J)^+$. Donc (cf. formule (13)),

$$\omega(e_I)\,\omega(e_J) = e_{I_+} e_{J_+} = (-1)^{\tau(I^+, J^+)} e_{I_+ \triangle J_+} =$$
$$= (-1)^{\tau(I, J)} e_{(I \triangle J)_+} = \omega(e_I e_J)$$

quels que soient les éléments de base e_I et e_J de l'algèbre $\mathrm{Cl}(n-1)$. Donc, ω est un homomorphisme d'algèbres et, par suite, un isomorphisme (puisqu'il établit une correspondance biunivoque entre leurs bases). □

A noter que l'isomorphisme réciproque ω^{-1} agit à l'aide de la formule

$$\omega^{-1}(u + ve_n) = u + v,$$

où, à gauche, u et v sont des éléments de l'algèbre $\mathrm{Cl}(n)$ (resp. de $\mathrm{Cl}^0(n)$ et $\mathrm{Cl}^1(n)$) ne contenant pas e_n, c'est-à-dire se développant suivant les vecteurs de base e_I avec $I \subset [n-1]$, et, à droite, les « mêmes » éléments de l'algèbre $\mathrm{Cl}(n-1)$.

Remarque 4. Nous savons que les algèbres $\mathrm{Cl}^0(n)$ et $\mathrm{Cl}^0_+(n)$ sont isomorphes. Donc, *l'algèbre* $\mathrm{Cl}(n-1)$ *est aussi isomorphe à l'algèbre* $\mathrm{Cl}^0_+(n)$. L'isomorphisme $\omega_+ : \mathrm{Cl}(n-1) \to \mathrm{Cl}^0_+(n)$ est visiblement défini par les formules

$$\omega_+(e_I) = \begin{cases} \bar{e}_I & \text{si } |I| \text{ est pair,} \\ e_n\bar{e}_I & \text{si } |I| \text{ est impair.} \end{cases}$$

□

D'après la proposition 8, on peut admettre que $\mathrm{Spin}(n) \subset \subset \mathrm{Cl}(n-1)$. Donc, en appliquant encore le théorème 2, on obtient *une représentation des éléments du groupe* Spin(n) *par des matrices d'ordre* $2^{\alpha(n-1)}$ *sur l'algèbre* $\mathbb{K}$, *où*

$$\mathbb{K} = \begin{cases} \mathbb{R} \ \textit{pour} \ n = 8m-1,\ 8m,\ 8m+1, \\ \mathbb{C} \ \textit{pour} \ n = 8m-2,\ 8m+2, \\ \mathbb{H} \ \textit{pour} \ n = 8m+3,\ 8m+4,\ 8m+5. \end{cases}$$

Certes, pour $n = 8m$ et $n = 8m+4$, on a affaire à une représentation par des couples de matrices d'ordre $2^{\alpha(n-1)-1}$.

Tout vecteur $u \in \mathbb{R}^n$ peut être mis sous la forme $u = u' + \lambda e_n$, où $u' \in \mathbb{R}^{n-1}$ et $\lambda \in \mathbb{R}$; de plus $u \in S^{n-1}$ si et seulement si $|u'|^2 + \lambda^2 = 1$. Par définition, le groupe Spin(n) est engendré par les éléments uv, où $u, v \in S^{n-1}$, c'est-à-dire par les éléments de la forme

$$(u' + \lambda e_n)(v' + \mu e_n) = (u'v' - \lambda\mu) + (\mu u' - \lambda v')e_n,$$

et, par suite, son image dans $\mathrm{Cl}(n-1)$ est engendrée par les éléments de la forme $(u'v' - \lambda\mu) + (\mu u' - \lambda v')$. Mais en posant $v^* = v' - \mu e_n$, on trouve que l'élément

$$ue_n \cdot v^* e_n = (u'e_n - \lambda)(v'e_n + \mu) = (u'v' - \lambda\mu) + (\mu u' - \lambda v')e_n$$

admet dans $\mathrm{Cl}^0(n-1)$ la même image $(u'v' - \lambda\mu) + (\mu u' - \lambda v')$ que l'élément uv. Ceci montre (puisque $ue_n \in \mathrm{Spin}(n)$ et $v^* \in S^{n-1}$ si $v \in S^{n-1}$) que l'image du groupe Spin(n) dans l'algèbre $\mathrm{Cl}^0(n-1)$ est engendrée par les images des éléments de la forme ue_n, $u \in S^{n-1}$. Comme $ue_n = u'e_n - \lambda$, ces images sont de la forme $u' - \lambda$ et, par suite, se représentent par des matrices de la forme $U - \lambda E$. Comme $\bar{U}^\top = -U$, et donc,

$$(U - \lambda E)(\bar{U} - \lambda E)^\top = -(U - \lambda E)(U + \lambda E) = \\ = -U^2 + \lambda^2 E = (|u'|^2 + \lambda^2)E = E,$$

toutes ces matrices sont orthogonales (resp. unitaires, symplectiques). Ceci prouve que *les représentations nouvellement construites sont aussi des représentations dans les groupes* $\mathrm{O}_{\mathbb{K}}(N)$ *(et dans les groupes* $\mathrm{SO}(N)$ *pour* $\mathbb{K} = \mathbb{R}$).

Donc, nous avons représenté maintenant le groupe Spin($8m+1$) par des matrices orthogonales d'ordre 2^{4m}, tandis qu'avant il l'était par des matrices unitaires du même ordre. Mais, en étudiant attentivement tous les isomorphismes, on démontrerait que *ces matrices unitaires sont en fait réelles* et que la représentation par des matrices réelles unitaires, i.e. orthogonales, est équivalente à celle qui vient juste d'être construite. (Ceci résulte sans calcul des résultats généraux sur les représentations matricielles des groupes Spin(n) qui seront prouvés dans le cadre de la théorie générale des représentations des groupes de Lie compacts. Nous glisserons donc sur ce fait ici.) Donc, nous n'avons pratiquement qu'une seule représentation du groupe Spin($8m+1$) par des matrices orthogonales, appelée *représentation spinorielle.*

Pour le groupe Spin($8m$), on obtient maintenant une représentation par des couples (A, B) de matrices orthogonales unimodulaires d'ordre 2^{4m-1}. Il s'avère (et ceci se démontre plus facilement dans le cadre de la théorie générale des représentations des groupes de Lie compacts) qu'en identifiant ces couples aux matrices (19) d'ordre 2^{4m}, on est conduit à une représentation équivalente à celle construite précédemment. Donc, ici aussi on n'obtient qu'une légère précision de la représentation précédente.

Du reste, on préfère généralement considérer non pas des couples de matrices, mais les composantes de ces couples séparément, c'est-à-dire deux homomorphismes

$$\text{Spin}(8m) \rightrightarrows \text{SO}(2^{4m-1}). \tag{20}$$

Ces homomorphismes s'appellent *représentations semi-spinorielles* du groupe Spin($8m$).

Signalons que *les représentations semi-spinorielles ne sont pas des monomorphismes*. En effet, l'isomorphisme $\text{Cl}_\varepsilon(n) \approx \text{Cl}_{-\varepsilon}(n-2) \otimes \otimes \text{Cl}_\varepsilon(2)$ de la proposition 7 envoie l'élément $e_{[n]} = e_1 \dots e_n$ de l'algèbre $\text{Cl}_\varepsilon(n)$ dans l'élément $e_{[n-2]} \otimes (e_1e_2)^{n-1} = (-1)^{n-1}e_{[n-2]} \otimes 1$ de l'algèbre $\text{Cl}_{-\varepsilon}(n-2) \otimes \text{Cl}_\varepsilon(2)$. Donc, l'isomorphisme itéré $\text{Cl}(n) \approx \text{Cl}(n-4) \otimes \mathbb{H}(2)$ (on se borne au cas $\varepsilon = -1$) transforme cet élément en l'élément $e_{[n-4]} \otimes E$, où E est la matrice unité. Une récurrence évidente montre maintenant que pour $n = 8m-1$, à l'élément $e_{[n]}$ est associé l'élément $e_{[3]} \otimes E$ de l'algèbre $\text{Cl}(3) \otimes \otimes \mathbb{H}(2^{4m-3})$, donc, l'élément $e_1 \otimes 1 \otimes E \approx (E, -E)$ de l'algèbre $\text{Cl}_+(1) \otimes \text{Cl}(2) \otimes \mathbb{H}(2^{4m-3}) \approx (\mathbb{R} \oplus \mathbb{R}) \otimes \mathbb{R}(2^{4m-1})$. Ceci prouve que l'élément $e_{[8m-1]}$ est envoyé dans la matrice unité E par une représentation semi-spinorielle (20), et l'élément $-e_{[8m-1]}$ par l'autre représentation (20). (On admet ici que $\text{Spin}(8m) \subset \subset \text{Cl}(8m-1)$; l'élément $e_{[8m-1]}$ est remplacé par l'élément $e_{[8m]}$ par l'inclusion canonique $\text{Spin}(8m) \subset \text{Cl}^0(8m)$.) □

Remarque 5. On démontre que les éléments $e_{[8m]}$ et $-e_{[8m]}$ sont les seuls éléments non triviaux du groupe Spin($8m$) à être envoyés dans la matrice unité par les homomorphismes (20).

L'ancienne représentation « par couples » du groupe Spin($8m-1$) peut aussi être décomposée en deux homomorphismes. Mais ces homomorphismes seront des représentations exactes (c'est-à-dire des monomorphismes) et seront équivalents à l'unique nouvelle représentation du groupe Spin ($8m-1$) dans le groupe SO(2^{4m-1}).

On obtient maintenant une représentation complexe du groupe Spin($8m-2$) dans le groupe U(2^{4m-2}). Cette représentation se transforme par l'inclusion $U(2^{4m-2}) \subset SO(2^{4m-1})$ en une représentation équivalente à l'ancienne représentation du groupe SO(2^{4m-1}).

L'ancienne représentation complexe du groupe Spin($8m-3$) se déduit de la même façon de la nouvelle représentation quaternionique.

L'ancienne représentation quaternionique du groupe Spin($8m+2$) est en fait une représentation complexe équivalente à la nouvelle représentation. Cette situation est identique à celle qui a prévalu pour $n = 8m + 1$.

L'ancienne représentation quaternionique « par couples » du groupe Spin($8m+3$) est composée de deux représentations équivalentes à la nouvelle.

L'ancienne représentation quaternionique du groupe Spin($8m+4$) est en fait une représentation « par blocs » qui se ramène à la nouvelle représentation « par couples ».

Les deux dernières situations sont identiques à celles qui se présentent pour les groupes Spin($8m-1$) et Spin($8m$) respectivement.

Comme déjà signalé, ces assertions résultent toutes automatiquement de la théorie générale. Leur vérification directe, quoique fastidieuse, est du domaine du possible. Nous la laissons à l'initiative du lecteur.

Achevons cette leçon par la démonstration de la formule (18) d'algèbre linéaire que nous avons adoptée sans la prouver. Pour la matrice de l'opérateur A, cette formule équivaut à une relation entre ses mineurs d'ordre deux et peut donc être prouvée par des calculs, assez laborieux sans doute, mais directs. Néanmoins pour dévoiler la teneur de cette formule, nous préférons en donner une démonstration plus conceptuelle sous une forme générale naturelle.

Soit $\mathscr{V}$ un espace vectoriel de dimension finie sur un corps $\mathbb{K}$, et soit comme toujours $\mathscr{V}'$ son dual. Soient par ailleurs $\Lambda^p(\mathscr{V})$ et $\Lambda^p(\mathscr{V}')$ les espaces des fonctionnelles p-linéaires antisymétriques sur les espaces $\mathscr{V}$ et $\mathscr{V}'$ respectivement; cf. II, 9. Pour tout p-vecteur $x_1 \wedge \ldots \wedge x_p \in \Lambda^p(\mathscr{V})$ et tout p-covecteur $\xi^1 \wedge \ldots \wedge \xi^p \in \Lambda^p(\mathscr{V}')$, posons

$$\langle x_1 \wedge \ldots \wedge x_p, \xi^1 \wedge \ldots \wedge \xi^p \rangle = \det | \xi^i(x_j) |_{i,j=1 \ldots p}.$$

On vérifie immédiatement que la fonction $\langle\,,\,\rangle$ se prolonge d'une seule façon par linéarité en des éléments quelconques des espaces vectoriels $\Lambda^p(\mathscr{V})$ et $\Lambda^p(\mathscr{V}')$ et est un couplage de ces deux espaces (cf. II, 4). Donc, *les espaces* $\Lambda^p(\mathscr{V})$ *et* $\Lambda^p(\mathscr{V}')$ *sont canoniquement duaux l'un de l'autre* et, par suite, l'espace $\Lambda^p(\mathscr{V}') = \Lambda^p(\mathscr{V})$ peut être identifié au dual $\Lambda^p(\mathscr{V})'$ de $\Lambda^p(\mathscr{V})$.

Pour toute fonctionnelle antisymétrique $y \in \Lambda^q(\mathscr{V})$ (pour la commodité, nous modifierons légèrement les notations adoptées dans II) la correspondance $x \mapsto x \wedge y$, où $x \in \Lambda^p(\mathscr{V})$ est une application linéaire $\Lambda^p(\mathscr{V}) \to \Lambda^{p+q}(\mathscr{V})$ et, par suite, définit l'application adjointe $\Lambda^{p+q}(\mathscr{V}') \to \Lambda^p(\mathscr{V}')$. L'image d'une fonctionnelle $z \in \Lambda^{p+q}(\mathscr{V}')$ par cette application est désignée par $y \lrcorner z$ et s'appelle *produit intérieur à gauche* des fonctionnelles y et z. Par définition

$$\langle x,\ y \lrcorner z\rangle = \langle x \wedge y,\ z\rangle \text{ pour tout } z \in \Lambda^{p+q}(\mathscr{V}').$$

De façon analogue, le *produit intérieur à droite* $x \llcorner y$ des fonctionnelles $x \in \Lambda^{p+q}(\mathscr{V})$ et $y \in \Lambda^q(\mathscr{V}')$, appartient à $\Lambda^p(\mathscr{V})$ et est caractérisé par l'égalité

$$\langle x \llcorner y,\ z\rangle = \langle x,\ y \wedge z\rangle \text{ pour tout } z \in \Lambda^p(\mathscr{V}').$$

Si $e_1, \ldots, e_n$ est une base de l'espace $\mathscr{V}$, alors pour tout sous-ensemble $I = \{i_1, \ldots, i_p\}$ de l'ensemble $[n] = \{1, \ldots, n\}$, où $i_1 < \ldots < i_p$, on désignera le p-vecteur de base $e_{i_1} \wedge \ldots \wedge e_{i_p}$ par e_I (voir ci-dessus les notations analogues pour les éléments de l'algèbre de Clifford). De façon analogue, on désignera par e^I le p-covecteur de base $e^{i_1} \wedge \ldots \wedge e^{i_p}$, où $e^1, \ldots, e^n$ sont les vecteurs de la base duale de l'espace $\mathscr{V}'$. Ceci étant, de l'anticommutativité du produit extérieur, il s'ensuit immédiatement pour tous sous-ensembles $I,\ J \subset [n]$

$$e_I \wedge e_J = \begin{cases} (-1)^{\tau(I,\,J)} e_{I \cup J}, & \text{si } I \cap J = \varnothing, \\ 0 & \text{dans le cas contraire,} \end{cases}$$

où $\tau(I, J)$ désigne comme dans la formule (13) le nombre de couples (i, j) tels que $i \in I$, $j \in J$ et $i \geqslant j$. Cette formule est valable de toute évidence pour le multicovecteur $e^I \wedge e^J$.

De ces formules il résulte aussitôt que

$$e_I \lrcorner e^J = \begin{cases} (-1)^{\tau(K,\,I)} e^K, & \text{si } I \subset J, \\ 0 & \text{dans le cas contraire,} \end{cases}$$

et de façon analogue

$$e_J \llcorner e_I = \begin{cases} (-1)^{\tau(J,\,K)} e^K, & \text{si } I \subset J, \\ 0 & \text{dans le cas contraire,} \end{cases}$$

où $K = J \setminus I$ est le complémentaire de I relativement à J.

Considérons maintenant un opérateur linéaire $A\colon \mathcal{V} \to \mathcal{V}$. Si pour tout multivecteur $x_1 \wedge \ldots \wedge x_p \in \Lambda^p(\mathcal{V})$, on pose

$$A_{[p]}(x_1 \wedge \ldots \wedge x_p) = Ax_1 \wedge \ldots \wedge Ax_p,$$

et qu'on prolonge $A_{[p]}$ par linéarité en un élément quelconque de l'espace $\Lambda^p(\mathcal{V})$, on définit de toute évidence de façon unique un opérateur linéaire $A_{[p]}\colon \Lambda^p(\mathcal{V}) \to \Lambda^p(\mathcal{V})$ appelé généralement *puissance extérieure p-ième* de l'opérateur A. (L'opérateur $\hat{A}$ envisagé plus haut n'est autre que l'opérateur $A_{[2]}$.)

Il est clair que

$$A_{[p+q]}(x \wedge y) = A_{[p]}x \wedge A_{[q]}y$$

quelles que soient les fonctionnelles $x \in \Lambda^p(\mathcal{V})$ et $y \in \Lambda^q(\mathcal{V})$. Par ailleurs, il résulte des définitions que

$$(A')_{[p]} = (A_{[p]})',$$

où le symbole « prime » désigne comme toujours (cf. II, 14) l'opérateur adjoint. Donc, la notation $A'_{[p]}$ est tout à fait correcte.

La clef de la formule (10) est donnée par le lemme suivant:

Lemme 4. *Pour toutes fonctionnelles $x \in \Lambda^q(\mathcal{V})$, $y \in \Lambda^{p+q}(\mathcal{V}')$ et tout opérateur $A\colon \mathcal{V} \to \mathcal{V}$, on a la relation*

$$x \lrcorner A'_{[p+q]}y = A'_{[p]}(A_{[q]}x \lrcorner y).$$

Démonstration. Pour toute fonctionnelle $z \in \Lambda^p(\mathcal{V})$, on a

$$\begin{aligned}\langle z, x \lrcorner A'_{[p+q]}y\rangle &= \langle z \wedge x, A'_{[p+q]}y\rangle = \\ &= \langle A_{[p+q]}(z \wedge x), y\rangle = \\ &= \langle A_{[p]}z \wedge A_{[q]}x, y\rangle = \\ &= \langle A_{[p]}z, A_{[q]}x \lrcorner y\rangle = \langle z, A'_{[p]}(A_{[q]}x \lrcorner y)\rangle. \quad \square\end{aligned}$$

Appliquons le lemme 4 au cas où $p + q = n$ et $y = e^{[n]} = e^1 \wedge \ldots \wedge e^n$. Comme de toute évidence

$$A'_{[n]}e^{[n]} = (\det A)\, e^{[n]},$$

d'après le lemme 3, pour toute fonctionnelle $x \in \Lambda^q(\mathcal{V})$ et tout opérateur $A\colon \mathcal{V} \to \mathcal{V}$, on a la formule

$$(\det A)(x \lrcorner e^{[n]}) = A'_{[p]}(A_{[q]}x \lrcorner e^{[n]}).$$

En posant

$$Tx = x \lrcorner e_{[n]}, \quad x \in \Lambda^q(\mathcal{V}),$$

on peut mettre cette formule sous la forme

$$(\det A)T = A'_{[p]} \circ T \circ A_{[q]}, \quad \text{où} \quad p + q = n.$$

L'expression explicite de l'application T est

$$Te_I = (-1)^{\tau(I)}e^J.$$

où $J = [n] \setminus I$ est le complémentaire de I relativement à $[n]$, et $\tau(I) = \tau(J, I)$ le nombre de couples (j, i) tels que $i \in I$, $j \notin I$ et $j \geqslant i$. On voit, en particulier, que T est un isomorphisme $\Lambda^q(\mathcal{V}) \to \Lambda^{n-q}(\mathcal{V}')$.

Dans le cas particulier où $\det A = 1$, on obtient la formule

$$A^c_{[p]} \circ T = T \circ A_{[q]}, \tag{21}$$

où $A^c = (A')^{-1}$.

Jusqu'ici toutes nos constructions étaient complètement intrinsèques (même l'isomorphisme T ne dépend de la base qu'à un facteur multiplicatif près, plus exactement, au déterminant près de la matrice de changement de base). Fixons maintenant une base $e_1, \ldots, e_n$ (c'est-à-dire en fait passons de l'espace vectoriel $\mathcal{V}$ à l'espace vectoriel $\mathbb{K}^n$) et identifions les espaces $\mathcal{V}$ et $\mathcal{V}'$ en égalant les coordonnées dans les bases $e_1, \ldots, e_n$ et $e^1, \ldots, e^n$ (en termes intrinsèques cela revient à définir sur l'espace vectoriel un produit scalaire, c'est-à-dire un couplage de cet espace avec lui-même). Alors, T sera un isomorphisme $\Lambda^q(\mathcal{V}) \to \Lambda^{n-q}(\mathcal{V})$ et sera défini par la formule

$$Te_I = (-1)^{\tau(I)}e_J$$

et, par suite, pour $n = 4$, $p = q = 2$, sera confondu avec l'isomorphisme T défini par les formules (17) (généralisées au cas d'un corps de base $\mathbb{K}$ quelconque). Donc, la formule (18) est un cas particulier de la formule (21) aux notations près.

On peut donc considérer que la formule (18) est entièrement prouvée.

Remarque 6. En mettant la formule générale (21) sous forme matricielle, on obtient une formule qui exprime au moyen des mineurs d'ordre $q = n - p$ d'une matrice non dégénérée les mineurs d'ordre p de la matrice inverse. La démonstration numérique de cette formule implique des calculs excessivement compliqués.

LEÇON 14

Doublage des algèbres.— Algèbres métriques.— Algèbres normées.— Automorphismes et dérivations des algèbres métriques.— Dérivations d'une algèbre doublée.— Dérivations et automorphismes de l'algèbre $\mathbb{H}$.— Algèbre des octaves.— Algèbre de Lie $\mathfrak{g}_2$.— Constantes de structure de l'algèbre de Lie $\mathfrak{g}_2^{\mathbb{C}}$.— Donnée de l'algèbre $\mathfrak{g}_2^{\mathbb{C}}$ par des générateurs et des relations.

La construction des quaternions avec les nombres complexes est entièrement calquée sur celle des nombres complexes avec les réels. Plus, ces deux constructions sont des cas particuliers d'une même construction générale.

Soit $\mathcal{A}$ une algèbre (comme toujours de dimension finie) sur le corps des réels $\mathbb{R}$, sur laquelle est définie la *conjugaison*, c'est-à-dire un antiautomorphisme involutif $a \mapsto \bar{a}$ (le cas $\bar{a} = a$, $\forall a \in \mathcal{A}$ n'est pas exclu).

Soit $\mathcal{A}^2$ l'espace vectoriel, somme directe de deux exemplaires d'un espace vectoriel $\mathcal{A}$, c'est-à-dire l'espace des couples (a, b), où $a, b \in \mathcal{A}$. Munissons $\mathcal{A}^2$ de la multiplication suivante :

$$(a, b)(u, v) = (au - \bar{v}b, b\bar{u} + va).$$

Une vérification immédiate montre que $\mathcal{A}^2$ est une algèbre (de dimension $2n$, où $n = \dim \mathcal{A}$) pour cette multiplication. Nous appellerons cette algèbre *doublage* de l'algèbre $\mathcal{A}$.

Il est évident que la correspondance $a \mapsto (a, 0)$ est un monomorphisme de l'algèbre $\mathcal{A}$ dans l'algèbre $\mathcal{A}^2$. Nous identifierons les éléments a et $(a, 0)$ et admettrons de ce fait que l'algèbre $\mathcal{A}$ est une sous-algèbre de $\mathcal{A}^2$. Si l'algèbre $\mathcal{A}$ est unitaire, l'élément $1 = (1, 0)$ sera aussi l'unité de l'algèbre $\mathcal{A}^2$.

Soit $e = (0, 1)$. Alors, $be = (0, b)$ et, par suite $(a, b) = a + be$ quels que soient $a, b \in \mathcal{A}$. Donc, tout élément de l'algèbre $\mathcal{A}^2$ se

représente d'une seule façon sous la forme $a + be$. Ceci étant,

(1)
$$a(be) = (ba)e, \quad (ae)b = (a\bar{b})e,$$
$$(ae)(be) = -\bar{b}a,$$

ce qui, combiné à la condition de distributivité, définit de façon unique la multiplication sur $\mathcal{A}^2$. En particulier, $e^2 = -1$.

Le doublage $\mathbb{R}^2$ du corps $\mathbb{R}$ est l'algèbre $\mathbb{C}$ des nombres complexes, et le doublage $\mathbb{C}^2$ de l'algèbre $\mathbb{C}$, l'algèbre $\mathbb{H}$ des quaternions. La première assertion est évidente. Pour établir la seconde, il faut expliciter les nombres complexes a et b de chaque élément $\xi = a + be$ de l'algèbre $\mathbb{C}^2$, c'est-à-dire poser $a = a_0 + a_1 i$, $b = a_2 + a_3 i$, et désigner e par j et ie par k. On obtient pour ξ la forme quaternionique usuelle

$$\xi = a_0 + a_1 i + a_2 j + a_3 k.$$

On vérifie automatiquement que les éléments i, j et $k = ij$ satisfont les identités quaternioniques classiques.

Etant donné que pour tout élément $a \in \mathcal{A}$, on a $ea = \bar{a}e$, l'algèbre $\mathcal{A}^2$, tout comme l'algèbre $\mathbb{H}$, est visiblement non commutative si la conjugaison définie sur $\mathcal{A}$ n'est pas une application identique. De même, puisque $a(be) = (ba)e$, l'algèbre $\mathcal{A}^2$, n'est pas associative si l'algèbre $\mathcal{A}$ n'est pas commutative. En particulier, *le doublage $\mathbb{H}^2$ de $\mathbb{H}$ n'est pas associatif.*

On voit donc que le doublage itératif détériore progressivement les propriétés algébriques de la multiplication.

Pour construire les doublages itérés il faut certes définir la conjugaison sur l'algèbre $\mathcal{A}^2$. Nous le ferons à l'aide de la formule

$$\overline{a + be} = \bar{a} - be,$$

qui pour $\mathcal{A} = \mathbb{R}$ se transforme en la conjugaison complexe ordinaire (si l'on remplace bien sûr e par i). Il est évident que cette application est involutive et linéaire. Un calcul immédiat montre que c'est un antiautomorphisme. Pour $\mathcal{A} = \mathbb{C}$, c'est la conjugaison habituelle de l'algèbre $\mathbb{H}$.

On dira qu'une algèbre unitaire $\mathcal{A}$ sur le corps $\mathbb{R}$ est une *algèbre métrique* si elle est munie d'une conjugaison $a \mapsto \bar{a}$ telle que pour tout $a \in \mathcal{A}$, l'élément $a\bar{a}$ appartient à $\mathbb{R}$ (plus exactement, à $\mathbb{R}\cdot 1$) et $a\bar{a} > 0$ pour $a \neq 0$. Le réel $|a| = \sqrt{a\bar{a}}$ sera appelé *norme* de a. Par définition $|a| = 0$ si et seulement si $a = 0$.

Une vérification immédiate montre que dans toute algèbre métrique $\mathcal{A}$ la formule

$$(x, y) = \frac{x\bar{y} + y\bar{x}}{2}$$

définit un produit scalaire. Donc, ce produit fait de toute algèbre métrique un espace euclidien, la norme $|a|$ de $a \in \mathcal{A}$ n'étant autre que sa longueur.

Désignons par $\mathcal{A}'$, l'orthocomplément de l'unité dans une algèbre métrique $\mathcal{A}$.

Tout élément $a \in \mathcal{A}$ se représente d'une seule manière sous la forme $a = \lambda + a'$, où $\lambda \in \mathbb{R}$ et $a' \in \mathcal{A}'$. Ceci étant, $\bar{a} = \lambda - a'$, de sorte qu'en particulier *$a \in \mathcal{A}'$ si et seulement si $\bar{a} = -a$, et $a \in \mathbb{R}$ si et seulement si $\bar{a} = a$.*

Par définition

$$x\bar{y} + \bar{x}y = 2(x, y) \tag{2}$$

quels que soient x, y de $\mathcal{A}$. En particulier, *si $x, y \in \mathcal{A}'$, alors $yx = -xy$ si et seulement si $x \perp y$.*

Il est aisé de voir que *le doublage $\mathcal{A}^2$ d'une algèbre métrique $\mathcal{A}$ est lui aussi métrique*. En effet,

$$(a + be)\overline{(a + be)} = (a + be)(\bar{a} - be) = a\bar{a} + b\bar{b}$$

quel que soit $a + be \in \mathcal{A}^2$. □

Sur $\mathcal{A}^2$ le produit scalaire est défini de toute évidence par la formule

$$(a + be, u + ve) = (a, u) + (b, v),$$

de sorte que la somme directe $\mathcal{A}^2 = \mathcal{A} \oplus \mathcal{A}$ est somme directe d'espaces euclidiens.

Toutes les algèbres $\mathbb{R}$, $\mathbb{C} = \mathbb{R}^2$, $\mathbb{H} = \mathbb{C}^2$, $\mathbb{H}^2$, ... sont donc des algèbres métriques.

Une algèbre $\mathcal{A}$ (mais à priori pas forcément métrique) de dimension finie munie d'une structure d'espace euclidien est une *algèbre normée* si

$$|ab| = |a| \cdot |b|$$

quels que soient les éléments a, $b \in \mathcal{A}$. Pour tout élément $a \neq 0$ d'une telle algèbre, les applications $x \mapsto \frac{ax}{|a|}$ et $x \mapsto \frac{xa}{|a|}$ sont des isométries et, par suite (puisque $\mathcal{A}$ est de dimension finie), des bijections. Donc, les équations $ax = b$ et $xa = b$ admettent une solution unique pour tout élément $b \in \mathcal{A}$, autrement dit *l'algèbre normée $\mathcal{A}$ est une algèbre à division.*

Les algèbres métriques $\mathbb{R}$, $\mathbb{C}$ et $\mathbb{H}$ sont des algèbres normées. Nous verrons ultérieurement que l'algèbre métrique $\mathbb{H}^2$ est aussi normée.

Le théorème de Hurwitz qui sera démontré au prochain semestre affirme que ces quatre algèbres sont les seules algèbres normées (de sorte que toute algèbre normée est nécessairement algèbre métrique).

Si un automorphisme $\Phi: \mathcal{A} \to \mathcal{A}$ d'une algèbre métrique commute à la conjugaison (c'est-à-dire si $\overline{\Phi a} = \Phi \bar{a}$ pour tout élément $a \in \mathcal{A}$), il est visiblement opérateur orthogonal. Réciproquement, si un automorphisme $\Phi: \mathcal{A} \to \mathcal{A}$ d'une algèbre métrique $\mathcal{A}$ est orthogonal, alors, puisque $\Phi 1 = 1$, il envoie dans lui-même le sous-espace $\mathcal{A}'$ et, par suite, commute à la conjugaison.

Cela étant, il est aisé de voir que *si une algèbre $\mathcal{A}$ est normée, tout automorphisme $\Phi: \mathcal{A} \to \mathcal{A}$ est un opérateur orthogonal.* En effet, il suffit de montrer que (cf. II, 21, condition b) de la proposition 2) si $|a| = 1$, alors $|\Phi a| = 1$. Mais si $|\Phi a| < 1$, alors $|\Phi a^k| = |(\Phi a)^k| = |\Phi a|^k \to 0$ pour $k \to \infty$, c'est-à-dire que $\Phi a^k \to 0$ et, par suite, $a^k \to 0$. Donc, $|a|^k = |a^k| \to 0$, ce qui est impossible pour $|a| = 1$. De façon analogue. si $|\Phi a| > 1$, alors $|a|^k \to \infty$ pour $k \to \infty$, ce qui est impossible aussi. Donc, $|\Phi a| = 1$. □

En principe, on fixera une base dans chacune des algèbres $\mathcal{A}$ étudiées. On peut donc identifier le groupe des opérateurs orthogonaux $\mathcal{A} \to \mathcal{A}$ au groupe $O(n)$ des matrices orthogonales. Avec cette identification, le groupe Aut $\mathcal{A}$ des automorphismes de l'algèbre normée $\mathcal{A}$ est justiciable de l'inclusion

$$\mathrm{Aut}\,\mathcal{A} \subset O(n), \quad \text{où} \quad n = \dim \mathcal{A}.$$

Bien plus, puisque tout automorphisme $\Phi: \mathcal{A} \to \mathcal{A}$ est défini de façon unique par l'application linéaire $\Phi': \mathcal{A}' \to \mathcal{A}'$ qu'il induit (si $a = \lambda + a'$, où $\lambda \in \mathbb{R}$, et $a' \in \mathcal{A}'$, alors $\Phi a = \lambda + \Phi' a'$), on peut, en identifiant Φ à Φ', admettre que

$$\mathrm{Aut}\,\mathcal{A} \subset O(n-1). \tag{3}$$

Ceci vaut, en particulier, pour $\mathcal{A} = \mathbb{C}$ lorsque $n = 2$. Donc, puisque $O(1) = \mathbb{Z}_2$, *le groupe* Aut $\mathbb{C}$ *est le groupe du second ordre* $\mathbb{Z}_2$ *composé de l'automorphisme identique* id *et de la conjugaison complexe* $a \mapsto \bar{a}$.

Pour l'algèbre de Lie Der $\mathcal{A} = \mathfrak{l}(\mathrm{Aut}\,\mathcal{A})$ des dérivations de l'algèbre $\mathcal{A}$, il s'ensuit de l'inclusion (3) que

$$\mathrm{Der}\,\mathcal{A} \subset \mathfrak{so}(n-1), \tag{4}$$

où $\mathfrak{so}(n-1)$ est l'algèbre de Lie des matrices antisymétriques d'ordre $n - 1$. Donc, en particulier, Der $\mathbb{C} = 0$.

On peut sans peine établir directement les égalités $\mathrm{Aut}\,\mathbb{C} = \mathbb{Z}_2$ et $\mathrm{Der}\,\mathbb{C} = 0$. En effet, étant un opérateur linéaire sur $\mathbb{R}$ envoyant 1 dans 1, tout automorphisme $\Phi: \mathbb{C} \to \mathbb{C}$ est défini de façon unique par le nombre Φi. Mais comme $i^2 = -1$, ce nombre doit vérifier l'égalité $(\Phi i)^2 = -1$. Donc, $\Phi i = \pm i$, ce qui nous donne l'opérateur identique et la conjugaison complexe.

De façon analogue, toute dérivation $D: \mathbb{C} \to \mathbb{C}$ est définie de façon unique par le nombre Di qui doit vérifier l'identité

$$Di \cdot i + i \cdot Di = D(i^2) = -D1 = 0$$

qui n'est possible que pour $Di = 0$.

Ces considérations élémentaires se généralisent de manière évidente au cas de toute algèbre doublée $\mathcal{A}^2$. Plus exactement, toute dérivation D de l'algèbre $\mathcal{A}^2$ définit à l'aide des formules

$$Da = D_0 a + Fa \cdot e, \quad De = x_0 + y_0 e$$

deux opérateurs linéaires D_0, F: $\mathcal{A} \to \mathcal{A}$ et deux éléments $x_0, y_0 \in \mathcal{A}$, et de plus, puisque

$$D(a + be) = Da + Db \cdot e + b \cdot De,$$

c'est-à-dire que

$$D(a + be) = (D_0 a - Fb + bx_0) + (Fa + D_0 b + y_0 b)e, \tag{5}$$

la dérivation D est définie de façon unique par les opérateurs D_0 et F et par les éléments x_0, y_0. Donc, pour décrire l'algèbre de Lie Der $\mathcal{A}^2$, il suffit de décrire les quadruplets (D_0, F, x_0, y_0) pour lesquels la formule (5) définit la dérivation de l'algèbre $\mathcal{A}^2$. (On pourrait donner une description analogue des automorphismes, mais les calculs seraient trop compliqués.)

Pour déterminer les conditions que doivent remplir D_0, F, x_0, y_0 pour que $D \in$ Der $\mathcal{A}^2$, il faut dans la relation

$$D(\xi\eta) = D\xi \cdot \eta + \xi \cdot D\eta, \tag{6}$$

où $\xi = a + be$ et $\eta = x + ye$ sont des éléments quelconques de $\mathcal{A}^2$, exprimer D à l'aide de la formule (5), effectuer tous les produits et identifier les coefficients en 1 et e. Par exemple, pour $\xi = a$, $\eta = x$, on obtient l'identité

$$D_0(ax) + F(ax)e = (D_0 a \cdot x + a \cdot D_0 x) + (Fa \cdot \bar{x} + Fx \cdot a)\, e,$$

d'où il s'ensuit que D_0 est la dérivation de l'algèbre $\mathcal{A}$, et F vérifie l'identité

$$F(ax) = Fa \cdot \bar{x} + Fx \cdot a. \tag{7}$$

On obtient des identités analogues pour $\xi = be$, $\eta = x$, pour $\xi = a$, $\eta = ye$ et pour $\xi = be$, $\eta = ye$. En pratique, il est plus payant de trouver d'abord la solution générale de l'équation fonctionnelle (7) et ensuite de prendre en considération ces identités. Par ailleurs, il est utile de tenir compte dès le départ des relations $x_0 + \bar{x}_0 = 0$ $y_0 + \bar{y}_0 = 0$, qui sont établies pour $\xi = \eta = e$ et qui expriment que $x_0, y_0 \in \mathcal{A}'$.

Supposons par exemple que $\mathcal{A} = \mathbb{C}$, donc que $\mathcal{A}^2 = \mathbb{H}$. Comme Der $\mathbb{C} = 0$, on a $D_0 = 0$. S'agissant de l'opérateur F, en utilisant le fait que $\mathbb{C} = \mathbb{R}^2$, on peut pour le calculer appliquer le même procédé en introduisant les opérateurs linéaires $P, Q: \mathbb{R} \to \mathbb{R}$ définis par la formule

$$Fx = Px + Qx \cdot i, \quad x \in \mathbb{R}.$$

Pour $a, x \in \mathbb{R}$, l'identité (7) entraîne

$$P(ax) = Pa \cdot x + Px \cdot a \quad \text{et} \quad Q(ax) = Qa \cdot x + Qx \cdot a,$$

c'est-à-dire que P et Q sont des dérivations du corps $\mathbb{R}$. Donc $P = Q = 0$, c'est-à-dire que $Fx = 0$ pour $x \in \mathbb{R}$. Par conséquent, $F(x + yi) = F(yi) = Fi \cdot y$, c'est-à-dire que

$$Fa = z_0 \cdot \operatorname{Im} a, \quad \text{où} \quad z_0 = Fi.$$

Par une vérification immédiate on s'assure que la relation (7) est satisfaite pour tout z_0. Ceci prouve que toute dérivation $D: \mathbb{H} \to \mathbb{H}$ doit être de la forme (cf. formule (5))

$$D(a + bj) = (-z_0 \operatorname{Im} b + bx_0) + (z_0 \operatorname{Im} a + y_0 b)\, j, \tag{8}$$

où $x_0, y_0, z_0 \in \mathbb{C}$. Ceci étant, les paramètres x_0 et y_0 doivent appartenir à $\mathbb{C}'$, c'est-à-dire être de la forme $x_0 = ia_0$, $y_0 = ib_0$, où $a_0, b_0 \in \mathbb{R}$. En portant (8) dans la formule générale (6), on constate aussitôt que l'application $D: \mathbb{H} \to \mathbb{H}$ définie par (8) est une dérivation de l'algèbre $\mathbb{H}$ si et seulement si $\operatorname{Re} z_0 = -a_0$, c'est-à-dire si $z_0 = -a_0 + ic_0$, où $c_0 \in \mathbb{R}$. Ceci prouve que *toute dérivation $D: \mathbb{H} \to \mathbb{H}$ de l'algèbre $\mathbb{H}$ est définie par la formule*

$$\begin{aligned} D(a + bj) = {} & i(a_0 \operatorname{Re} b - c_0 \operatorname{Im} b) + \\ & + [-(a_0 \operatorname{Im} a + b_0 \operatorname{Im} b) + i(b_0 \operatorname{Re} b + c_0 \operatorname{Im} a)]\, j, \end{aligned} \tag{9}$$

où a_0, b_0, c_0 sont des réels arbitraires.

L'application (9) envoie $\mathbb{H}'$ dans $\mathbb{H}'$ et dans une base i, j, k de $\mathbb{H}'$ est définie par la matrice

$$\begin{pmatrix} 0 & a_0 & -c_0 \\ -a_0 & 0 & -b_0 \\ c_0 & b_0 & 0 \end{pmatrix}.$$

Donc, Der $\mathbb{H} \subset \mathfrak{so}(3)$ conformément à la formule générale (4). Bien plus, on voit maintenant que

$$\operatorname{Der} \mathbb{H} = \mathfrak{so}(3).$$

De cette égalité il résulte que le groupe Aut $\mathbb{H}$ est confondu soit avec le groupe SO(3), soit avec le groupe O(3) (puisque Aut $\mathbb{H} \subset O(3)$ en vertu de la formule générale (3)). Mais comme l'appli-

cation définie par la matrice $-E$ n'est visiblement pas un automorphisme, il vient

$$\text{Aut}\,\mathbb{H} = SO(3). \tag{10}$$

On remarque, en particulier, que, contrairement au groupe Aut $\mathbb{C}$, *le groupe* Aut $\mathbb{H}$ *est connexe.*

D'autre part, en comparant l'égalité (10) avec la proposition 4 de la leçon précédente, on trouve immédiatement que *tout automorphisme* $\mathbb{H} \to \mathbb{H}$ *est un automorphisme intérieur de la forme* $\eta \mapsto \xi\eta\xi^{-1}$, *où* $\xi \in S^3$.

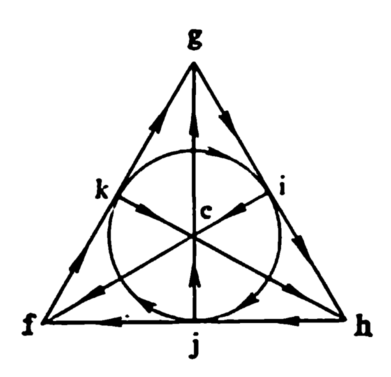
Remarque 1. En fait on peut prouver l'égalité (10) sans calcul en utilisant la proposition 4 de la leçon précédente. En effet, cette proposition dit que le groupe des automorphismes intérieurs $\eta \mapsto \xi\eta\xi^{-1}$, $\xi \in S^3$. de l'algèbre $\mathbb{H}$ est confondu avec le groupe SO(3). Donc, Aut $\mathbb{H} \supset$ SO (3). D'autre part nous venons tout juste d'établir que l'égalité Aut $\mathbb{H}$ = O(3) est impossible. Donc Aut $\mathbb{H}$ = = SO (3). Nous avons produit une démonstration directe, car elle nous servira de modèle pour le cas plus compliqué de l'algèbre $\mathbb{H}^2$.

L'algèbre $\mathbb{H}^2$ est appelée *algèbre des octaves* (ou *algèbre des nombres de Cayley*) et ses éléments, *octaves* (ou *nombres de Cayley*). En l'honneur de Cayley, on la désigne par $\mathbb{C}$a bien que ce fut Graves et non Cayley qui la découvrit le premier.

Par définition, chaque octave est de la forme $\xi = a + be$, où a et b sont des quaternions. Les octaves se multiplient à l'aide des formules (1). Une base de l'algèbre $\mathbb{C}$a est composée de 1 et des 7 éléments

$$i, \quad j, \quad k, \quad e, \quad f = ie, \quad g = je, \quad h = ke, \tag{11}$$

dont le carré est égal à -1 et le produit deux à deux est représenté schématiquement sur la figure suivante:



Le produit de deux éléments quelconques de (11) est égal, au signe près, au troisième élément de la même droite (ou du cercle), le signe dépendant de l'orientation de cette droite. Par exemple, $eh = k$ et $fj = -h$.

Comme déjà signalé, l'algèbre Ca n'est pas associative. Mais *elle est alternative*, c'est-à-dire que pour deux éléments quelconques ξ, η on a les identités

$$(\xi\eta)\eta = \xi(\eta\eta), \quad \xi(\xi\eta) = (\xi\xi)\eta.$$

En effet, soient $\xi = a + be$, $\eta = u + ve$. Alors

$$\xi\eta = (au - \bar{v}b) + (b\bar{u} + va)e,$$

$$(\xi\eta)\eta = [(au - \bar{v}b)u - \bar{v}(b\bar{u} + va)] + \\ + [(b\bar{u} + va)\bar{u} + v(au - \bar{v}b)]e,$$

$$\eta\eta = (u^2 - \bar{v}v) + (v\bar{u} + vu)e,$$

$$\xi(\eta\eta) = [a(u^2 - \bar{v}v) - \overline{(v\bar{u} + vu)}b] + [b\overline{(u^2 - \bar{v}v)} + (v\bar{u} + vu)a]e;$$

comme les nombres $\bar{v}v = v\bar{v}$ et $u + \bar{u}$ sont réels, donc commutent à tout quaternion, il vient

$$\begin{aligned} a(u^2 - \bar{v}v) - \overline{(v\bar{u} + vu)}b &= au^2 - a\bar{v}v - (u + \bar{u})\bar{v}b = \\ &= au^2 - \bar{v}va - \bar{v}b(u + \bar{u}) = \\ &= (au - \bar{v}b)u - \bar{v}(b\bar{u} + va) \end{aligned}$$

et

$$\begin{aligned} b\overline{(u^2 - \bar{v}v)} + (v\bar{u} + vu)a &= b\bar{u}^2 - b\bar{v}v + v(\bar{u} + u)a = \\ &= b\bar{u}^2 - v\bar{v}b + va(\bar{u} + u) = \\ &= (b\bar{u} + va)\bar{u} + v(au - \bar{v}b). \end{aligned}$$

Donc, $(\xi\eta)\eta = \xi(\eta\eta)$.

L'égalité $\xi(\xi\eta) = (\xi\xi)\eta$ s'établit de façon analogue. □

D'après la théorie générale, la conjugaison sur Ca est définie par la formule

$$\overline{a + be} = \bar{a} - be.$$

Cette opération laisse invariant l'élément 1 et change le signe de chaque élément (11), de sorte que les éléments (11) forment une base du sous-espace Ca'.

E. Artin a montré que deux éléments quelconques d'une algèbre alternative engendrent une sous-algèbre associative. Donc, le même raisonnement appliqué aux quaternions montre que *l'algèbre des octaves* Ca *est normée.* Mais la démonstration du théorème d'Artin est assez laborieuse, nous la sauterons faute de temps. Nous prouverons donc par des calculs directs que Ca est normée.

Soient $\xi = a + be$ et $\eta = u + ve$ deux octaves. Il nous faut montrer que $|\xi\eta| = |\xi| \cdot |\eta|$. D'après ce qui précède

$$|\xi\eta|^2 = |au - \bar{v}b|^2 + |b\bar{u} + va|^2 =$$
$$= (au - \bar{v}b)(\bar{u}\,\bar{a} - \bar{b}v) + (b\bar{u} + va)(u\bar{b} + \bar{a}\,\bar{v})$$

et

$$|\xi|^2 |\eta|^2 = (a\bar{a} + b\bar{b})(u\bar{u} + v\bar{v}).$$

En posant donc $v = \lambda + v'$, où $\lambda \in \mathbb{R}$ et $v' \in \mathrm{Ca}'$ (donc $\bar{v'} = -v'$), on trouve que

$$|\xi\eta|^2 - |\xi|^2|\eta|^2 = \lambda\,(au\bar{b} + b\bar{u}\,\bar{a} - b\bar{u}a - au\bar{b}) +$$
$$+ (au\bar{b} + b\bar{u}\,\bar{a})\,v' - v'(\bar{a}\bar{u}\,b + b\bar{u}\,\bar{a}) = 0,$$

car $au\bar{b} + b\bar{u}\,\bar{a}$ est réel, donc commute au quaternion v'. □

De là il résulte, en particulier, que *l'algèbre* Ca *est munie d'une division*. Du reste, comme dans le cas des quaternions, on vérifie immédiatement (à l'aide de la propriété d'alternativité) que les équations $\xi x = \eta$ et $x\xi = \eta$ sont satisfaites (pour $\xi \neq 0$) respectivement par les octaves $x = \xi^{-1}\eta$ et $x = \eta\xi^{-1}$, où $\xi^{-1} = \frac{\bar{\xi}}{|\xi|^2}$. □

Remarque 2. L'itération du doublage fait apparaître les algèbres métriques Ca^2, $(\mathrm{Ca}^2)^2$, . . ., etc. Mais ces algèbres présentent un intérêt minime dans la mesure où elles ne sont ni normées (d'après le théorème de Hurwitz) ni alternatives (d'après le théorème généralisé de Frobenius, toute algèbre alternative de dimension finie sur le corps $\mathbb{R}$ est isomorphe à une algèbre $\mathbb{R}$, $\mathbb{C}$, **H** ou Ca; cf. par exemple [11]). Bien plus, ces algèbres ne sont même pas des algèbres à division, car on démontre (le premier à le faire a été F. Adams) par de puissantes méthodes de topologie algébrique qu'une algèbre à division sur le corps $\mathbb{R}$ ne peut être que de dimension 1, 2, 4 et 8 (de plus parmi les algèbres de dimension 4 et 8, il existe des algèbres à division autres que les algèbres **H** et Ca.

Pour des raisons dans le détail desquelles nous n'entrerons pas, le groupe des automorphismes Aut Ca de l'algèbre Ca est désigné par G_2, est son algèbre de Lie Der Ca, par $\mathfrak{g}_2$. L'algèbre Ca étant normée et dim Ca = 8, on a, en vertu de (3),

$$G_2 \subset \mathrm{O}(7)$$

et, par suite,

$$\mathfrak{g}_2 \subset \mathfrak{so}(7),$$

où $\mathfrak{so}(7)$ est l'algèbre de Lie des matrices antisymétriques d'ordre 7. Mais contrairement au cas précédent, l'algèbre $\mathfrak{g}_2$ n'est pas confon-

17—01642

due avec l'algèbre $\mathfrak{so}(7)$, de sorte que maintenant nous avons affaire à une nouvelle algèbre de Lie (et à un nouveau groupe de Lie) que nous n'avons jamais rencontrée précédemment. Nous prouverons cette non-coïncidence en montrant par le calcul que dim $\mathfrak{g}_2 = 14$, alors que dim $\mathfrak{so}(7) = 21$.

Le calcul de l'algèbre $\mathfrak{g}_2$ s'effectue par un procédé déjà connu. Toute dérivation $D \in \mathfrak{g}_2$ est définie par la formule (5), où désormais x_0 et y_0 sont des quaternions, D_0 une dérivation de l'algèbre $\mathbb{H}$, et F un opérateur linéaire $\mathbb{H} \to \mathbb{H}$ vérifiant d'identité (7). En posant $Fx = Px + Qx\cdot j$, où maintenant $x \in \mathbb{C}$ et P, Q sont des opérateurs linéaires $\mathbb{C} \to \mathbb{C}$, on obtient pour P et Q les équations

$$P(xy) = Px\cdot\bar{y} + Py\cdot x, \qquad Q(xy) = Qx\cdot y + Qy\cdot\bar{x},$$

dont la solution générale est de la forme

$$Px = a_0 \operatorname{Im} x, \quad Qx = b_0 \operatorname{Im} x,$$

où a_0, $b_0 \in \mathbb{C}$. Donc, en posant $Fj = z_0 + w_0 j$, où x_0, $w_0 \in \mathbb{C}$, on trouve que la solution générale de l'équation fonctionnelle (5) sur $\mathbb{H}$ est de la forme

$$(12) \quad F\xi = (a_0 \operatorname{Im} x + b_0 \operatorname{Im} y + z_0 y) + \\ + (b_0 \operatorname{Im} x - a_0 \operatorname{Im} y + w_0 \bar{y})\, j, \quad \xi = x + yj,$$

où a_0, b_0, z_0, $w_0 \in \mathbb{C}$. Un calcul automatique, quoique harassant, montre maintenant que l'application D: $\mathbb{C}\mathrm{a} \to \mathbb{C}\mathrm{a}$ définie par la formule (5), où D_0 est une dérivation de l'algèbre $\mathbb{H}$, F l'application (12), x_0 et y_0 des quaternions de $\mathbb{H}'$, est une dérivation de l'algèbre $\mathbb{C}\mathrm{a}$, c'est-à-dire appartient à $\mathfrak{g}_2$, si et seulement si $x_0 = -i\operatorname{Re} a_0 - (\bar{z}_0 + i \operatorname{Re} b_0)\, j$.

Nous avons ainsi prouvé le lemme suivant:

Lemme 1. *Toute dérivation D de l'algèbre des octaves $\mathbb{C}\mathrm{a}$ est définie par:*

a) *une dérivation quelconque D_0 de l'algèbre des quaternions $\mathbb{H}$;*

b) *un quaternion y_0 vérifiant la relation $\bar{y}_0 = -y_0$;*

c) *quatre nombres complexes a_0, b_0, z_0, w_0.*

Cette dérivation agit d'après la formule

$$D(\xi + \eta e) = (D_0\xi - F\eta + \eta x_0) + (F\xi + D_0\eta + y_0\eta)e,$$

où F est une application $\mathbb{H} \to \mathbb{H}$ définie par la formule (12), *et* $x_0 = -i \operatorname{Re} a_0 - (\bar{z}_0 + i \operatorname{Re} b_0)j$. □

La dérivation D_0 est définie par une matrice antisymétrique du troisième ordre, et, par suite, dépend de trois paramètres réels, le quaternion y_0, de trois aussi, et le quadruplet, a, b, x_0, y_0, de huit.

Donc, la dérivation D est définie par la donnée de 14 (=3 + 3 + 8) paramètres réels indépendants. Donc, dim $\mathfrak{g}_2 = 14$.

La matrice associée à la dérivation D dans la base (11) peut être représentée sous la forme conventionnelle mais suggestive suivante:

$$
\begin{array}{c|ccc|c|c|}
 & i & j & k & e & f \quad g \quad h \\
\hline
i & & & & -a_0 & -b_0 \\
j & & D_0 & & -z_0 & -w_0 \\
k & & & & -c_0 & d_0 \\
\hline
e,\ f & a_0 & z_0 & c_0 & 0 & -y_0 \\
\hline
g,\ h & b_0 & w_0 & -d_0 & y_0 & D_0 + y_0^{\bullet} \\
\hline
\end{array}
$$

où $c_0 = b_0 + z_0 i$, $d_0 = a_0 + w_0 i$, et $y_0^{\bullet}$ désigne un opérateur (de $\mathbf{H}'$ dans $\mathbf{H}$) de multiplication par y_0. (Cet opérateur est défini par la matrice

$$
\begin{pmatrix}
-b_1 & -b_2 & -b_3 \\
0 & -b_3 & b_2 \\
b_3 & 0 & -b_1 \\
-b_2 & b_1 & 0
\end{pmatrix},
$$

où $b_1 i + b_2 j + b_3 k = y_0$.) De là il s'ensuit que parmi les matrices antisymétriques (a_{ij}) d'ordre 7 les matrices de $\mathfrak{g}_2$ sont caractérisées par les sept conditions suivantes:

$$
\begin{gathered}
a_{32} + a_{45} + a_{76} = 0, \\
a_{13} + a_{64} + a_{75} = 0, \qquad a_{21} + a_{65} + a_{47} = 0, \\
a_{14} + a_{36} + a_{27} = 0, \qquad a_{51} + a_{26} + a_{73} = 0, \\
a_{17} + a_{42} + a_{53} = 0, \qquad a_{61} + a_{52} + a_{34} = 0.
\end{gathered}
\tag{13}
$$

Si $a_0 = b_0 = 0$ et

$$
D_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & p \\ 0 & -p & 0 \end{pmatrix},
$$

c'est-à-dire si $Di = 0$, alors la dérivation D enverra dans lui-même le sous-espace $\mathscr{V}$ de l'algèbre $\mathbb{C}a$ orthogonal aux éléments 1 et i. Ce sous-espace est muni de la multiplication par i qui en fait un espace vectoriel sur $\mathbb{C}$ admettant la base j, e, g. Un calcul immédiat montre que dans cette base la dérivation D est définie par la matrice complexe

$$\begin{pmatrix} -pi & -\bar{z}_0 & -v_0 \\ z_0 & b_1 i & -v_0 \\ \bar{v}_0 & \bar{v}_0 & (p-b_1)\, i \end{pmatrix},$$

où $v_0 = b_2 + b_3 i$, et p et b_1 sont des nombres réels. Comme ceci est exactement la forme des matrices antihermitiennes du troisième ordre de trace nulle, on obtient que *les dérivations* D: $\mathbb{C}a \to \mathbb{C}a$ *telles que* $Di = 0$ *forment une sous-algèbre de l'algèbre de Lie* $\mathfrak{g}_2$, *isomorphe à l'algèbre de Lie* $\mathfrak{su}(3)$.

Un procédé standard de description des algèbres de dimension finie consiste à donner leurs *constantes de structure* dans une certaine base, c'est-à-dire les coefficients des développements suivant une base des produits deux à deux des éléments de cette base. Donc, si $e_1, \ldots, e_n$ est une base de $\mathscr{A}$, ses constantes de structure c^k_{ij} sont définies par la formule

$$e_i e_j = c^k_{ij} e_k, \quad i, j, k = 1, \ldots, n.$$

Le nombre total de ces constantes est n^3.

Appliquée à l'algèbre de Lie $\mathfrak{g}_2$, cette méthode implique le choix de 2744 ($=14^3$) nombres, ce qui en pratique est irréel. La situation n'est pas améliorée par la prise en considération de l'antisymétrie des constantes par rapport à i et j, puisqu'il reste encore 1279 ($=91\cdot 14 = \frac{14\cdot 13}{2}\cdot 14$) constantes. Mais, on peut escompter que si la base est convenablement choisie, on pourra mettre en évidence des lois les décrivant de manière satisfaisante. Il s'avère que cela est possible et la situation est simplifiée à l'extrême par la complexification de l'algèbre de Lie $\mathfrak{g}_2$, c'est-à-dire pour l'algèbre de Lie $\mathfrak{g}_2^{\mathbb{C}} = \mathfrak{g}_2 \otimes \mathbb{C}$ sur le corps $\mathbb{C}$, composée des matrices complexes d'ordre 7 vérifiant les conditions (13). Comme le passage inverse de $\mathfrak{g}_2^{\mathbb{C}}$ à $\mathfrak{g}_2$ est aisé à réaliser, ceci nous fournit les constantes de structure de l'algèbre de Lie $\mathfrak{g}_2$.

Considérons la base canonique de l'algèbre de Lie $\mathfrak{so}(7)$, composée des matrices

$$E_{[i,j]} = \frac{E_{ij} - E_{ji}}{2},$$

où $i, j = 1, \ldots, 7$ et $i < j$. Des conditions (13) il s'ensuit aussitôt que les matrices

$$\begin{aligned} P_0 &= E_{[3,2]} + E_{[6,7]}, & Q_0 &= E_{[4,5]} + E_{[6,7]}, \\ P_1 &= E_{[1,3]} + E_{[5,7]}, & Q_1 &= E_{[6,4]} + E_{[5,7]}, \\ P_2 &= E_{[2,1]} + E_{[7,4]}, & Q_2 &= E_{[6,5]} + E_{[7,4]}, \\ P_3 &= E_{[1,4]} + E_{[7,2]}, & Q_3 &= E_{[3,6]} + E_{[7,2]}, \\ P_4 &= E_{[5,1]} + E_{[3,7]}, & Q_4 &= E_{[2,6]} + E_{[3,7]}, \\ P_5 &= E_{[1,7]} + E_{[3,5]}, & Q_5 &= E_{[4,2]} + E_{[3,5]}, \\ P_6 &= E_{[6,1]} + E_{[4,3]}, & Q_6 &= E_{[5,2]} + E_{[4,3]} \end{aligned}$$

appartiennent toutes à l'algèbre de Lie $\mathfrak{g}_2$. Ces matrices étant visiblement linéairement indépendantes, elles forment une base (sur $\mathbb{R}$) de l'algèbre de Lie $\mathfrak{g}_2$.

Appesantissons-nous sur les combinaisons linéaires

$$(14) \qquad H = aP_0 + bQ_0 = aE_{[3,2]} + bE_{[4,5]} + cE_{[7,6]}$$

des matrices P_0 et Q_0, où $a + b + c = 0$. On remarquera que $[H_1, H_2] = 0$ pour tous éléments H_1, H_2 de la forme (14).

Un calcul immédiat sur les matrices nous montre que

$$\begin{aligned} [H, P_1] &= aP_2 + cQ_2, & [H, Q_1] &= (c - b)\, Q_2, \\ [H, P_2] &= -aP_1 - cQ_1, & [H, Q_2] &= (b - c)\, Q_1, \\ [H, P_3] &= bP_4 + cQ_4, & [H, Q_3] &= (c - a)\, Q_4, \\ [H, P_4] &= -bP_3 - cQ_3, & [H, Q_4] &= (a - c)\, Q_3, \\ [H, P_5] &= cP_6 + aQ_6, & [H, Q_5] &= (a - b)\, Q_6, \\ [H, P_6] &= -cP_5 - aQ_5, & [H, Q_6] &= (b - a)\, Q_5, \end{aligned}$$

d'où il résulte aussitôt que les éléments

$$\begin{aligned} U_{\pm 1} &= (2P_2 - Q_2) \pm i\,(2P_1 - Q_2), & V_{\pm 1} &= Q_2 \pm iQ_1, \\ U_{\pm 2} &= (2P_4 - Q_4) \pm i\,(2P_3 - Q_3), & V_{\pm 2} &= Q_4 \pm iQ_3, \\ U_{\pm 3} &= (2P_6 - Q_6) \pm i\,(2P_5 - Q_5), & V_{\pm 3} &= Q_6 \pm iQ_5, \end{aligned}$$

qui forment la base de l'algèbre complexifiée $\mathfrak{g}_2^{\mathbf{C}}$ satisfont les relations

$$(15) \qquad \begin{aligned} [H, U_{\pm 1}] &= \pm iaU_{\pm 1}, & [H, V_{\pm 1}] &= \pm i\,(c - b)\, V_{\pm 1}, \\ [H, U_{\pm 2}] &= \pm ibU_{\pm 1}, & [H, V_{\pm 2}] &= \pm i\,(c - a)\, V_{\pm 2}, \\ [H, U_{\pm 3}] &= \pm icU_{\pm 3}, & [H, V_{\pm 3}] &= \pm i\,(a - b)\, V_{\pm 3}. \end{aligned}$$

Pour écrire ces relations sous une forme plus compendieuse, il vaut la peine d'introduire l'espace réel $\mathfrak{h}^*$ à deux dimensions dual

de l'espace $\mathfrak{h}$ des éléments (14). Soient e_1, e_2 la base de $\mathfrak{h}^*$ adjointe d'une base P_0, Q_0 de $\mathfrak{h}$. Alors, pour tout élément (14) de $\mathfrak{h}$, on aura les formules

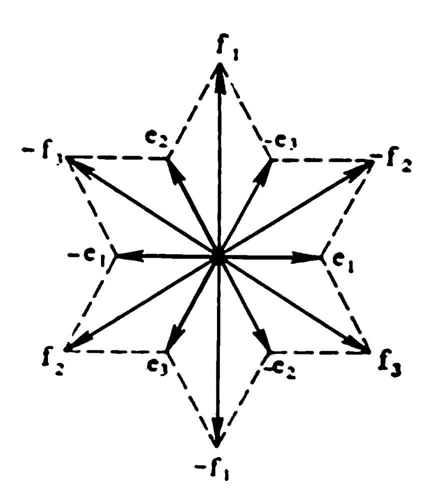
(16) $$e_1(H) = a, \quad e_2(H) = b, \quad e_3(H) = c,$$

où $e_3 = -(e_1 + e_2)$.

Il est commode (mais pas obligatoire) d'admettre que l'espace $\mathfrak{h}^*$ est muni d'une structure euclidienne et est rapporté à un système de coordonnées rectangulaires dans lequel les composantes des vecteurs e_1 et e_2 sont respectivement $\left(\frac{\sqrt{6}}{3}, 0\right)$ et $\left(-\frac{\sqrt{6}}{6}, \frac{\sqrt{2}}{2}\right)$. Alors les vecteurs $\pm e_1$, $\pm e_2$, $\pm e_3$ et les vecteurs $\pm f_1$, $\pm f_2$, $\pm f_3$, où

$$f_1 = e_2 - e_3, \quad f_2 = e_3 - e_1, \quad f_3 = e_1 - e_2$$

seront les rayons vecteurs des sommets d'un dodécagone étoilé régulier. L'ensemble de ces douze vecteurs du plan sera appelé *configuration* G_2.



A tout vecteur $\alpha \in G_2$ associons maintenant un élément X_α de l'algèbre de Lie $\mathfrak{g}_2^{\mathbb{C}}$ en posant

$$X_\alpha = \begin{cases} U_{\pm k} & \text{si } \alpha = \pm e_k, \\ V_{\pm k} & \text{si } \alpha = \pm f_k. \end{cases}$$

En vertu des relations (16), les formules (15) peuvent alors être regroupées en une seule:

(17) $$[H, X_\alpha] = i\alpha(H)\, X_\alpha.$$

Les éléments X_α, $\alpha \in G_2$, forment avec tout couple H_1, H_2 d'éléments linéairement indépendants de $\mathfrak{h}$ une base de $\mathfrak{g}_2^{\mathbb{C}}$. Comme par construction $\overline{X}_\alpha = X_{-\alpha}$, les éléments H_1, H_2 et $X_\alpha + X_{-\alpha}$, $\frac{1}{i}(X_\alpha - X_{-\alpha})$ formeront une base de $\mathfrak{g}_2$. Etant donné que les

constantes de structure de la dernière base s'expriment de manière évidente en fonction de celles de la base H_1, H_2, X_α, il nous suffit de trouver simplement ces dernières. Les crochets $[H_1, X_\alpha]$ et $[H_2, X_\alpha]$ se calculant directement avec la formule (17) et le crochet $[H_1, H_2]$ étant nul, tout se ramène donc au calcul des crochets $[X_\alpha, X_\beta]$ pour tous les couples (α, β) non ordonnés des vecteurs de la configuration G_2 (ces couples sont au nombre de 66).

Il est aisé de voir que tout vecteur $\alpha \in \mathfrak{h}^*$ se représente d'une seule façon par $\alpha = ae_1 + be_2 + ce_3$, où $a + b + c = 0$, et de plus l'isomorphisme $\mathfrak{h}^* \approx \mathfrak{h}$ induit par la structure euclidienne dont nous avons muni $\mathfrak{h}^*$ envoie ce vecteur précisément dans l'élément (14) de l'espace $\mathfrak{h}$. Nous pouvons donc désigner aussi l'élément (14) par α. On désignera par H_α le produit de α par $\frac{2}{|\alpha|^2}$. L'expression de H_α est

$$H_\alpha = \frac{2}{a^2+b^2+c^2}(aE_{[3,2]} + bE_{[4,5]} + cE_{[7,6]}),$$

puisque, ce qui est immédiat, $|\alpha|^2 = a^2 + b^2 + c^2$.

Proposition 1. *Pour tous vecteurs* α, $\beta \in G_2$, *on a*

(18) $$[X_\alpha, X_{-\alpha}] = iH_\alpha,$$

(19) $$[X_\alpha, X_\beta] = 0 \textit{ pour } \beta \neq -\alpha \textit{ et } \alpha + \beta \notin G_2,$$

(20) $$[X_\alpha, X_\beta] = N_{\alpha,\beta}X_{\alpha+\beta} \textit{ pour } \alpha + \beta \in G_2 \ (\textit{donc } \beta \neq -\alpha).$$

Ici $N_{\alpha,\beta}$ *sont des entiers dont les modules satisfont la relation*

(21) $$|N_{\alpha,\beta}| = p + 1,$$

où p *est le plus grand des nombres entiers tels que le vecteur* $\beta - j\alpha$ *appartienne à la configuration* G_2 *pour tout* $j = 0, 1, \ldots, p$. □

Cette proposition sera prouvée dans le prochain semestre à l'aide d'une théorie générale. Tout ce que nous pouvons proposer au lecteur, c'est de vérifier ceci méthodiquement par des calculs dans les 66 cas.

Remarque 3. Ces calculs nous donneront évidemment des coefficients $N_{\alpha,\beta}$ signés (en tout 30 couples (α, β) pour lesquels $N_{\alpha,\beta} \neq 0$). On pourrait certes énoncer une règle pour déterminer ces signes, mais, *primo*, elle est assez compliquée, et, *secundo*, elle n'est pas intrinsèque. C'est que les relations (17), (18) et l'égalité $\overline{X}_\alpha = X_{-\alpha}$ caractérisent les éléments X_α au signe près. En ce sens les éléments $\pm X_\alpha$ sont définis intrinsèquement et sans ambiguïté dans l'algèbre $\mathfrak{g}_2^{\mathbf{C}}$. Le choix des signes dont dépendent les signes des coefficients $N_{\alpha,\beta}$ ne peut être décrit par aucun procédé intrinsèque.

Compte tenu de cette remarque on peut dire que la proposition 1 définit complètement les constantes de structure de l'algèbre de Lie $\mathfrak{g}_2^{\mathbf{C}}$ (donc, de l'algèbre de Lie $\mathfrak{g}_2$).

Les algèbres peuvent être définies aussi par leurs générateurs et par des relations.

Proposition 2. *L'algèbre de Lie $\mathfrak{g}_2^{\mathbb{C}}$ peut être donnée par quatre générateurs X_1, Y_1, X_2, Y_2 satisfaisant les quinze relations suivantes*

$$
(22)\qquad
\begin{gathered}
[[X_1, Y_1], [X_2, Y_2]] = 0,\\
[X_1, Y_2] = 0, \quad [X_2, Y_1] = 0,\\
[[X_1, Y_1], X_1] - 2X_1 = 0, \ [[X_1, Y_1], X_2] + X_2 = 0,\\
[[X_2, Y_2], X_1] + 3X_1 = 0, \ [[X_2, Y_2], X_2] - 2X_2 = 0,\\
[[X_1, Y_1], Y_1] + 2Y_1 = 0, \ [[X_1, Y_1], Y_2] - Y_2 = 0,\\
[[X_2, Y_2], Y_1] - 3Y_1 = 0, \ [[X_2, Y_2], Y_2] + 2Y_2 = 0,\\
[X_1 [X_1, X_2]] = 0, \ [Y_1, [Y_1, Y_2]] = 0,\\
[X_2, [X_2 [X_2 [X_2, X_1]]]] = 0,\\
[Y_2, [Y_2, [Y_2, [Y_2, Y_1]]]] = 0.
\end{gathered}
$$

A la démonstration formelle de la proposition 2 nous préférons une discussion détaillée des relations (22).

En convenant que $H_1 = -i [X_1, Y_1]$, $H_2 = -i [X_2, Y_2]$ et en introduisant les nombres

$$n_{11} = n_{22} = 2, \quad n_{12} = -1, \quad n_{21} = -3,$$

on peut mettre les relations (22) sous la forme plus concise suivante

$$
(23)\qquad
\begin{gathered}
[H_p, H_q] = 0,\\
[X_p, Y_q] = 0 \quad \text{pour} \quad p \neq q,\\
[H_p, X_q] = in_{p,q}X_q, \quad [H_p, Y_q] = -in_{p,q}Y_q,\\
(\text{ad } X_p)^{|n_{pq}|+1} X_q = 0, \ (\text{ad } Y_p)^{|n_{pq}|+1} Y_q = 0.
\end{gathered}
$$

C'est sous cette forme que nous les utiliserons.

Il est aisé de voir que *les relations* (23) *sont satisfaites dans l'algèbre $\mathfrak{g}_2^{\mathbb{C}}$ pour* $X_1 = X_{f_1}$, $Y_1 = X_{-f_1}$, $X_2 = X_{e_2}$, $Y_2 = X_{-e_2}$ (*et* $H_1 = H_{-f_1}$, $H_2 = H_{e_2}$). En effet, la première relation est vérifiée, car, comme déjà signalé, le crochet de Lie est nul pour tous éléments de $\mathfrak{h}$. La deuxième relation résulte de la formule (19), puisque $f_1 - e_2 = e_3 - 2e_2 \notin G_2$ et $e_2 - f_1 = 2e_2 - e_3 \notin G_2$. Les troisième et quatrième relations sont des cas particuliers de la formule (17), puisque

$$\alpha(H_\beta) = \left(\alpha, \frac{2\beta}{|\beta|^2}\right) - \frac{2(\alpha, \beta)}{|\beta|^2} = \begin{cases} 2 \text{ pour } \alpha = \beta = f_1, \ e_2, \\ -3 \text{ pour } \alpha = f_1, \ \beta = e_2, \\ -1 \text{ pour } \alpha = e_2, \ \beta = f_1 \end{cases}$$

(on se sert ici de l'identification $\mathfrak{h} = \mathfrak{h}^*$), quant aux deux dernières, elles découlent des formules (19) et (20) de la proposition 1, car

$$\begin{aligned} &f_1 + e_2 = e_3 \in G_2, \text{ mais } f_1 + e_3 \notin G_2, \\ (24) \qquad &e_2 + f_1 = e_3 \in G_2, \quad e_2 + e_3 = -e_1 \in G_2, \\ &e_2 - e_1 = -f_3 \in G_2, \text{ mais } e_2 - f_3 \notin G_2. \quad \square \end{aligned}$$

Si l'on introduit maintenant une algèbre de Lie $\mathfrak{l}$ libre de générateurs X_1, X_2, Y_1, Y_2 et son homomorphisme φ dans l'algèbre $\mathfrak{g}_2^{\mathbb{C}}$, défini par les formules

$$\varphi(X_1) = X_{f_1}, \quad \varphi(X_2) = X_{e_2},$$
$$\varphi(Y_1) = X_{-f_1}, \quad \varphi(Y_2) = X_{-e_2},$$

alors l'assertion prouvée exprime que l'homomorphisme φ annule les premiers membres de toutes les relations (22), et, donc, induit un homomorphisme $\bar{\varphi}$: $\bar{\mathfrak{l}} \to \mathfrak{g}_2^{\mathbb{C}}$ dans l'algèbre de Lie $\mathfrak{g}_2^{\mathbb{C}}$ de l'algèbre quotient $\bar{\mathfrak{l}}$ de l'algèbre de Lie $\mathfrak{l}$ par l'idéal engendré par ces premiers membres. La proposition 2 revient maintenant à affirmer que *l'homomorphisme $\bar{\varphi}$ est un isomorphisme.* C'est sous cette forme que nous la prouverons.

Pour ne pas introduire de notations supplémentaires, on désignera les éléments de l'algèbre $\bar{\mathfrak{l}}$ par les mêmes symboles que les éléments correspondants de l'algèbre $\mathfrak{l}$. Le symbole $X \sim Y$ désignera la proportionnalité des éléments X et Y de $\bar{\mathfrak{l}}$.

Dans l'algèbre $\bar{\mathfrak{l}}$ considérons les éléments

$$\begin{aligned} &H_1 = [X_1, Y_1], \quad H_2 = [X_2, Y_2], \\ &X_1, X_2, X_3 = [X_2, X_1], \quad X_4 = [X_2, X_3], \\ (25) \qquad &X_5 = [X_2, X_4], \quad X_6 = [X_4, X_3], \\ &Y_1, Y_2, Y_3 = [Y_2, Y_1], \quad Y_4 = [Y_2, Y_3], \\ &Y_5 = [Y_2, Y_4], \quad Y_6 = [Y_4, Y_3]. \end{aligned}$$

Il s'avère que *pour tous éléments U, V de la liste* (25), *l'élément* $[U, V]$ *est soit nul, soit proportionnel à un élément de cette liste.* En effet, par hypothèse $[H_1, H_2] = 0$, $[H_i, X_1] \sim X_1$, $[H_i, X_2] \sim X_2$. $i = 1, 2$. Or si $[H, U] \sim U$ et $[H, V] \sim V$, alors en vertu de l'identité de Jacobi

$$[H, [U, V]] = [[H, U], V] + [U, [H, V]] \sim [U, V].$$

Donc, par une récurrence évidente, on obtient

$$[H_p, X_q] \sim X_q$$

pour tout q. On établit de façon analogue que

$$[H_p, Y_q] \sim Y_q.$$

D'autre part, par hypothèse $[X_1, X_2] = X_3$ et $[X_1, X_3] = [X_1, [X_1, X_2]] = 0$. Donc, $[X_1, X_4] = [[X_1, X_2], X_3] + [X_2, [X_1, X_3]] = 0$, $[X_1, X_5] = [X_3, X_4] = -X_6$ et $[X_1, X_6] = 0$. Comme $[X_1, Y_1] = H_1$, $[X_1, Y_2] = 0$, alors $[X_1, Y_3] = [[X_1, Y_2], Y_1] + [Y_2, [X_1, Y_1]] = -[[X_1, Y_1], Y_2] \sim [H_1, Y_2] \sim Y_2$, et, par suite, $[X_1, Y_4] = -[[X_1, Y_3], Y_2] = 0$, $[X_1, Y_5] = 0$, $[X_1, Y_6] \sim [[X_1, Y_3], Y_4] \sim [Y_2, Y_4] = Y_5$. Les autres crochets se calculent de la même manière. □

De la proposition prouvée il résulte que l'enveloppe linéaire des éléments (25) est une sous-algèbre de l'algèbre de Lie $\bar{\mathfrak{l}}$ et, par suite — puisqu'elle contient les générateurs X_1, X_2 et Y_1, Y_2 — est confondue avec cette algèbre. Ceci prouve, en particulier, que $\dim \bar{\mathfrak{l}} \leqslant 14$.

Nous pouvons maintenant prouver sans peine la proposition 2.

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 2. Des résultats établis ci-dessus (cf., en particulier, les formules (24)), il s'ensuit aussitôt que l'homomorphisme $\bar{\varphi}$ envoie les éléments (25) dans des éléments proportionnels respectivement aux éléments

$$
\begin{gathered}
H_{f_1},\ H_{e_1},\\
X_{f_1},\ X_{e_2},\ X_{e_3},\ X_{-e_1},\ X_{-f_3},\ X_{f_2},\\
X_{-f_1},\ X_{-e_2},\ X_{-e_3},\ X_{e_1},\ X_{f_3},\ X_{-f_2}
\end{gathered}
\tag{26}
$$

(s'agissant des éléments X_6 et Y_6, ceci résulte de l'égalité $e_3 - e_1 = f_2$). Les éléments (26) formant une base de l'algèbre $\mathfrak{g}_2^{\mathbb{C}}$, l'homomorphisme $\bar{\varphi}_2$ est un épimorphisme et, par suite, un isomorphisme, puisque $\dim \bar{\mathfrak{l}} \leqslant 14 = \dim \mathfrak{g}_2^{\mathbb{C}}$. □

Nous avons ainsi étudié l'algèbre de Lie $\mathfrak{g}_2$ pratiquement sous tous les points de vue possibles.

LEÇON 15

Identités dans l'algèbre des octaves Ca.— Sous-algèbres de l'algèbre des octaves Ca.— Groupe de Lie G_2. — Principe de trialité pour le groupe Spin(8).— Analogue du principe de trialité pour le groupe Spin(9).— Algèbre d'Albert Al.— Plan projectif des octaves.

Maintenant qu'on a étudié l'algèbre de Lie $\mathfrak{g}_2$, on peut se tourner vers le groupe de Lie $G_2 = \text{Aut } \mathrm{Ca}$ correspondant. Les résultats de la leçon précédente sont valables au revêtement près pour la structure algébrique de la composante de l'unité du groupe G_2. Le problème majeur qui reste à élucider est de savoir si le groupe G_2 est connexe, et s'il l'est, de déterminer son groupe fondamental.

Ceci implique une étude plus poussée de la structure algébrique de l'algèbre Ca.

Soit $\mathcal{A}$ une algèbre alternative. En remplaçant b par $x + y$ dans la loi alternative $(ab)b = a(bb)$, en chassant les parenthèses et en réduisant les termes semblables, on obtient l'identité

$$ax \cdot y + ay \cdot x = a \cdot xy + a \cdot yx$$

(pour alléger les formules on écrit des points à la place des parenthèses). Ce procédé de déduction d'une identité à partir d'une autre s'appelle *polarisation* (ou *linéarisation*).

De façon analogue, en polarisant la deuxième loi alternative $b(ba) = (bb)a$, on obtient — après avoir remplacé b par $x + y$ et désigné x, y, a respectivement par a, x, y — l'identité

$$ax \cdot y + xa \cdot y = a \cdot xy + x \cdot ay.$$

La première identité exprime que l'expression $(ax)y - a(xy)$ (appelée *associateur* des éléments a, x, y) est antisymétrique en x et y, et la deuxième, qu'elle est antisymétrique en a et x. Mais alors cette expression est antisymétrique en a et y, ce qui nous donne la troi-

sième loi alternative

$$ax \cdot y + yx \cdot a = a \cdot xy + y \cdot xa,$$

qui pour $y = a$ devient $(ax)a = a(xa)$ (cette identité s'appelle aussi *identité d'élasticité*).

Si à l'instar de l'algèbre Ca, l'algèbre $\mathcal{A}$ est de plus métrique, alors pour tous éléments $a \in \mathcal{A}$ et $b = \lambda + b'$, où $\lambda \in \mathbb{R}$ et $b' \perp 1$, on aura l'égalité $(ab)(\lambda + b') = a \cdot b(\lambda + b')$, c'est-à-dire l'égalité $ab \cdot b' = a \cdot bb'$. Donc, $ab \cdot \lambda - ab \cdot b' = a \cdot b\lambda - a \cdot bb'$, c'est-à-dire que $ab \cdot \bar{b} = a \cdot b\bar{b} = a(b, b)$. La polarisation de cette identité nous donne l'identité

$$ax \cdot \bar{y} + ay \cdot \bar{x} = 2\,(x, y)a, \tag{1}$$

qui est une généralisation de l'identité (2) de la leçon précédente (pour $a = 1$).

Supposons maintenant que l'algèbre $\mathcal{A}$ est normée, c'est-à-dire qu'y est satisfaite l'identité $(ab, ab) = (a, a)(b, b)$. En polarisant cette identité d'abord sur $b = x + y$ et ensuite sur $a = u + v$, on obtient l'identité

$$(ux, vy) + (vx, uy) = 2\,(u, v)(x, y), \tag{2}$$

pour tous éléments u, v, x, y de $\mathcal{A}$ (et, par suite, de Ca).

Intéressons-nous tout particulièrement au sous-ensemble de l'espace vectoriel Ca$'$, composé des éléments ξ tels que $|\xi| = 1$. Ce sous-ensemble est une sphère à 6 dimensions que nous désignerons par S^6.

Il est immédiat de voir que $\xi \in S^6$ *si et seulement si* $\xi^2 = -1$. En effet, si $\xi \in$ Ca$'$, alors $\bar{\xi} = -\xi$, et, par suite, $\xi^2 = -|\xi|^2$. Donc si $|\xi| = 1$, alors $\xi^2 = -1$. Réciproquement, si $\xi^2 = -1$, alors $|\xi| = 1$, et, par suite, $\xi\bar{\xi} = 1$, c'est-à-dire que $\xi(-\bar{\xi}) = -1 = \xi\xi$. Par conséquent, $\bar{\xi} = -\xi$, c'est-à-dire que $\bar{\xi} \in$ Ca$'$, et comme $|\xi| = 1$, alors $\xi \in S^6$. □

Par linéarité il s'ensuit de là que $\xi \in$ Ca$'$ *si et seulement si* $\xi^2 = -|\xi|^2$.

Soit maintenant $\mathcal{H}$ une sous-algèbre unitaire (donc fermée pour la conjugaison) de l'algèbre Ca, distincte de Ca, et soit ζ une octave de S^6 orthogonale à $\mathcal{H}$. *Pour tout élément* $b \in \mathcal{H}$, *l'octave* $b\zeta$ *est* alors *orthogonale à la sous-algèbre* $\mathcal{H}$. En effet, en posant dans (2) $u = 1$, $v = b$, $x = \zeta$ et $y = a$, où $a \in \mathcal{H}$, et en tenant compte du fait que $(ab, \zeta) = 0$ (car $ab \in \mathcal{H}$) et $(\zeta, 1) = 0$ (par hypothèse), on obtient aussitôt $(b\zeta, a) = 0$ pour tout élément $a \in \mathcal{H}$. □

En particulier, $b\zeta \perp 1$ (puisque $1 \in \mathcal{H}$), de sorte que $\overline{b\zeta} = -b\zeta$.

Par ailleurs, *pour tous éléments* a, $b \in \mathcal{H}$, *on a les égalités*

$$a\cdot b\zeta = ba\cdot\zeta, \quad a\zeta\cdot b = a\bar{b}\cdot\zeta, \tag{3}$$
$$a\zeta\cdot b\zeta = -\bar{b}a.$$

En effet, en faisant $x = \zeta$, $y = \bar{b}$ et en tenant compte du fait que $\zeta \perp b$, donc que $\zeta \perp \bar{b}$, on obtient l'égalité

$$a\zeta\cdot b + a\bar{b}\cdot\bar{\zeta} = 0,$$

qui est équivalente à la deuxième identité (3). De façon analogue, en faisant $a = 1$, $x = a$, $y = \overline{b\zeta} = -b\zeta$ dans (1), on est conduit à l'égalité

$$-a\cdot b\zeta + b\zeta\cdot\bar{a} = -2\,(a,\, b\zeta) = 0$$

et, par suite, à l'égalité

$$a\cdot b\zeta = b\zeta\cdot\bar{a},$$

qui, d'après ce qui a été démontré, est équivalente à la première identité (3). Enfin pour $b \in \mathbb{R}$, la troisième identité (3) devient $a\zeta\cdot b\zeta = -ba$ et, par suite, se ramène à la deuxième identité. Donc, il suffit de prouver cette identité pour $b \perp 1$ seulement. Mais, dans ce cas, en posant $x = \zeta$ et $y = \overline{b\zeta} = -b\zeta$ dans (1), on obtient

$$a\zeta\cdot b\zeta + (a\cdot b\zeta)\zeta = -2\,(\zeta,\, b\zeta)a = 0,$$

puisque pour $u = 1$, $v = b$, $x = y = \zeta$ il s'ensuit de (2) que $(\zeta,\, b\zeta) = (1,\, b)\,(\zeta,\, \zeta) = 0$. Donc en vertu des identités déjà établies, $a\zeta\cdot b\zeta = -(ba)\zeta\cdot\zeta = ba = -\bar{b}a$. □

Il est aisé de voir maintenant que *l'algèbre* $\mathcal{H}$ *est associative*. En effet, si a, b, $c \in \mathcal{H}$, alors en remplaçant a, x, y respectivement par $b\zeta$, c et $\bar{a}\zeta$ dans (1), on obtient l'identité

$$(b\zeta\cdot\bar{c})\cdot\overline{\bar{a}\zeta} + (b\zeta\cdot\bar{a}\zeta)c = 2\,(\bar{c},\, \bar{a}\zeta)\cdot b\zeta = 0,$$

qui est équivalente, en vertu des identités (3), à la relation associative $(ab)c = a(bc)$. □

Nous pouvons maintenant entamer la démonstration du lemme fondamental des automorphismes de l'algèbre $\mathbb{C}a$.

Un automorphisme arbitraire $\Phi\colon \mathbb{C}a \to \mathbb{C}a$ transforme les éléments i, j et e en des éléments $\xi = \Phi i$, $\eta = \Phi j$ et $\zeta = \Phi e$ de S^6 tels que η est orthogonal à ξ, et ζ à ξ, η et $\xi\eta$. Il s'avère que ces conditions sont non seulement nécessaires, mais aussi suffisantes pour l'existence de l'automorphisme Φ:

Lemme 1. *Pour tous éléments* ξ, η, $\zeta \in S^6$ *tels que*

a) η *soit orthogonal à* ξ;

b) ζ *soit orthogonal à* ξ, η *et* $\xi\eta$,
il existe (visiblement un seul) automorphisme Φ *de l'algèbre* Ca *tel que*

$$\xi = \Phi i, \quad \eta = \Phi j, \quad \zeta = \Phi e.$$

D é m o n s t r a t i o n. Comme $\xi \in S^6$ et $\eta \in S^6$, il vient $\overline{\xi}^2 = -1$ et $\eta^2 = -1$ et puisque $\xi \perp \eta$, alors $\xi\eta = -\eta\xi$. Donc, $\overline{\xi\eta} = \overline{\eta}\overline{\xi} = \eta\xi = -\xi\eta$ et, par suite, $\xi\eta \in S^6$ (puisque $|\xi\eta| = |\xi| \times |\eta| = 1$). Donc, $(\xi\eta)^2 = -1$. Par ailleurs, en vertu de la loi alternative on a $\xi(\xi\eta) = -\eta$ et $(\xi\eta)\eta = -\xi$; en posant $u = v = \xi$, $x = \eta$ et $y = 1$ dans (28), on trouve que $(\xi\eta, \xi) = (\xi, \xi)(\eta, 1) = 0$, d'où il résulte que $(\xi\eta)\xi = -\xi(\xi\eta) = \eta$. Comme, de façon analogue, $\eta(\xi\eta) = \xi$, on voit qu'en multipliant ξ et η dans n'importe quel ordre et autant de fois qu'on le veut, on n'obtiendra que les éléments ± 1, $\pm\xi$, $\pm\eta$ et $\pm\xi\eta$. Ceci exprime que les éléments de la forme

$$a + b\xi + c\eta + d\xi\eta, \quad a, b, c, d \in \mathbb{R},$$

forment une sous-algèbre $\mathcal{H}$ de dimension 4 de l'algèbre Ca qui, d'après ce qui a été démontré plus haut, est une algèbre associative. Mais alors il est aisé de voir que les correspondances $1 \mapsto 1$, $i \mapsto \xi$, $j \mapsto \eta$, $k \rightarrow \xi\eta$ définissent un isomorphisme de l'algèbre des quaternions $\mathbb{H}$ sur l'algèbre $\mathcal{H}$.

Par ailleurs, l'élément ζ, orthogonal par hypothèse aux éléments 1, ξ, η et $\xi\eta$, est orthogonal à l'algèbre $\mathcal{H}$ tout entière. Il vérifie donc les identités (3). Mais, en comparant ces identités avec les formules (1) de la leçon 14, on constate aussitôt que l'isomorphisme $\mathbb{H} \rightarrow \mathcal{H}$ se prolonge en un homomorphisme (par lequel $e \mapsto \zeta$) de l'algèbre Ca sur la sous-algèbre engendrée par la sous-algèbre $\mathbb{H}$ et l'élément ζ. Etant donné que tout homomorphisme non nul d'une algèbre unitaire à division est nécessairement un monomorphisme (si l'image de $\xi \neq 0$ est 0, quelle est celle de ξ^{-1}?) et que tout monomorphisme d'une algèbre de dimension finie dans elle-même est nécessairement un automorphisme (car un opérateur injectif linéaire agissant sur un espace de dimension finie est bijectif), nous avons bien construit un automorphisme Ca $\rightarrow$ Ca envoyant i, j, e en ξ, η, ζ.

Ceci achève la démonstration du lemme 1. □

Du lemme 1, il résulte, en particulier, que *le groupe* $G_2 = \operatorname{Aut}$ Ca *opère transitivement sur la sphère* S^6, c'est-à-dire que l'application $G_2 \rightarrow S^6$ définie par la formule $\Phi \mapsto \Phi i$ est surjective. Cela signifie que la sphère S^6 est difféomorphe à la variété quotient G_2/K du groupe G_2 par le sous-groupe K des automorphismes laissant invariant l'élément i:

$$G_2/K \approx S^6.$$

Pour tout automorphisme $\Phi \in K$, l'élément $\eta = \Phi j$ de S^6 est orthogonal à l'élément i et, par suite, appartient à une sphère S^5 de

dimension 5 (l'équateur de la sphère S^6 de pôle i). Ceci étant, en vertu du lemme 1, l'application $\Phi \mapsto \Phi j$ du groupe K sur la sphère S^5 est surjective. Donc,

$$K/L \approx S^5,$$

où L est un sous-groupe du groupe K, composé des automorphismes laissant invariant l'élément j.

Mais pour les automorphismes Φ de L, l'élément $\zeta = \Phi e$ de S^6 est orthogonal aux éléments i, j, k, c'est-à-dire appartient à une sphère à trois dimensions $S^3 \subset S^6$. Ceci étant, en vertu du même lemme 1, l'application $\Phi \mapsto \Phi e$ est un difféomorphisme du groupe L sur la sphère S^3:

$$L \approx S^3.$$

Dans le langage topologique, cela signifie que le groupe G_2 est l'espace fibré de base S^6 et de fibres difféomorphes au groupe K, groupe qui à son tour est l'espace fibré de base S^5 et de fibres difféomorphes à S^3.

Proposition 1. *Le groupe G_2 est connexe et simplement connexe. Tout groupe de Lie localement isomorphe au groupe G_2 est isomorphe à G_2.*

D é m o n s t r a t i o n. La première assertion résulte immédiatement des résultats tout juste acquis en vertu du lemme 2 de la leçon 1, du corollaire du lemme 1 de la leçon 9 et de la proposition 7 de la leçon 12. La deuxième assertion signifie que le groupe G_2 n'admet pas de sous-groupes invariants discrets non triviaux. Comme (lemme 6 de la leçon 9) tous les sous-groupes invariants discrets d'un groupe de Lie connexe appartiennent à son centre, pour prouver cette assertion, il suffit de montrer que le centre du groupe G_2 est trivial, c'est-à-dire qu'un automorphisme Φ_0 de l'algèbre $\mathbb{C}a$ commutant à tout automorphisme Φ de $\mathbb{C}a$ est nécessairement un automorphisme identique. Mais si l'automorphisme Φ_0 commute à l'automorphisme Φ, et si $\Phi \in K$, c'est-à-dire que $\Phi i = i$, alors $\Phi(\Phi_0 i) = \Phi_0 i$. Or, ainsi qu'il résulte directement du lemme 1, la dernière égalité est valable pour tous les éléments du groupe K si seulement $\Phi_0 i = i$. On démontre de façon analogue que $\Phi_0 j = j$. Donc, $\Phi_0 = \text{id}$. □

Le groupe de Lie G_2 étant simplement connexe, sa structure algébrique est entièrement définie par celle de l'algèbre de Lie $\mathfrak{g}_2$.

Remarque 1. Comme déjà signalé dans la leçon précédente, le sous-espace $\mathcal{V}$ de l'algèbre $\mathbb{C}a$, orthogonal aux éléments 1 et i, est un espace vectoriel sur le corps $\mathbb{C}$ admettant la base j, e, g. Le produit scalaire sur $\mathbb{C}a$ induit sur $\mathcal{V}$ un produit scalaire hermitien pour lequel la base j, e, g est orthonormée. Tout automorphisme Φ: $\mathbb{C}a \to \mathbb{C}a$ laissant i invariant, c'est-à-dire appartenant au sous-groupe K, définit un opérateur $\mathcal{V} \to \mathcal{V}$ linéaire sur $\mathbb{C}$. Cet opérateur préserve le produit scalaire, c'est-à-dire est un opérateur uni-

taire. Comme son déterminant est de toute évidence égal à l'unité, le groupe K s'identifie à un sous-groupe du groupe SU(3). Ceci étant, du lemme 1 il résulte aussitôt que le groupe K est effectivement confondu avec le groupe SU(3). Donc, on peut considérer que $SU(3) \subset \subset G_2$, et de plus $G_2/SU(3) \approx S^6$.

L'algèbre de Lie $\mathfrak{k}$ du groupe K est composée des dérivations D telles que $Di = 0$. Nous avons ainsi encore prouvé que ces dérivations forment une algèbre de Lie isomorphe à l'algèbre de Lie $\mathfrak{su}(3)$.

L'algèbre Ca peut être appliquée à l'étude du groupe SO(8), puisque tout élément du groupe SO(8) peut être traité comme un opérateur orthogonal Ca $\to$ Ca. Du reste, il est plus commode de passer ici du groupe SO(8) à son revêtement universel, le groupe Spin(8).

Les résultats ultérieurs relatifs aux groupes Spin(8) et Spin(9) sont dus à N. Jacobson.

En ignorant la multiplication sur Ca, c'est-à-dire en traitant Ca simplement comme un espace euclidien, on peut construire les algèbres de Clifford $Cl_+(Ca)$ et $Cl(Ca')$. Donc, les éléments $i, j, k, e, f = ie, g = je, h=ke$ seront maintenant les générateurs de l'algèbre $Cl(Ca')$. Es qualités nous les désignerons, conformément à la leçon 13, par $e_1, \ldots, e_7$. Dans l'algèbre $Cl_+(Ca)$, il faut ajouter à ces générateurs l'unité de l'algèbre Ca. On se démarquera des notations adoptées dans la leçon 13 et l'on désignera ce générateur supplémentaire par e_0.

En vertu de la remarque qui suit la proposition 8 de la leçon 13, l'algèbre $Cl(Ca')$ est isomorphe à l'algèbre $Cl_+^0(Ca)$, de plus dans l'isomorphisme de cette remarque le rôle de l'élément e_n incombera dorénavant à l'élément e_0. Donc, cet isomorphisme sera défini par la formule

$$\omega_+(e_I) = \begin{cases} \bar{e}_I & \text{si } |I| \text{ est pair,} \\ e_0\bar{e}_I & \text{si } |I| \text{ est impair,} \end{cases}$$

où I est un sous-ensemble de l'ensemble $\{1, \ldots, 7\}$.

L'espace vectoriel Ca est par définition injecté dans l'algèbre $Cl_+(Ca)$. Nous admettrons qu'il est aussi injecté dans l'algèbre $Cl(Ca')$, en identifiant son unité avec l'unité 1 de l'algèbre $Cl(Ca')$. Il est aisé de voir qu'avec cette identification, pour tout élément de l'algèbre $Cl_+^0(Ca)$ de la forme e_0u, où $u \in Ca$, on aura

$$\omega_+^{-1}(e_0u) = u.$$

En effet, si $u = \lambda e_0 + u'$, où $\lambda \in \mathbb{R}$, et $u' \in Ca'$, alors $e_0u = \lambda + + e_0u'$ et, par suite, $\omega_+^{-1}(e_0u) = \lambda + \bar{u}' = \lambda + u' = u$ (la conjugaison ici est celle de $Cl_+(Ca)$ et non celle de Ca, donc $\bar{u}' = = u'$). □

L'algèbre Ca étant alternative, pour tous éléments $x \in \mathrm{Ca}$ et $u \in \mathrm{Ca}'$, on a l'égalité

$$(xu)\, u = x \cdot u^2 = - \mid u \mid^2 x$$

(on rappelle que $u^2 = - \mid u \mid^2$ si $u \in \mathrm{Ca}'$), qui exprime que pour l'opérateur linéaire R_u: $x \mapsto xu$, $u \in \mathrm{Ca}'$, on a

$$R_u^2 = - \mid u \mid^2 E.$$

Donc, la correspondance $u \to R_u$ se prolonge en un homomorphisme

$$R : \mathrm{Cl}\,(\mathrm{Ca}') \to \mathrm{End}\,\mathrm{Ca},$$

où End Ca est l'algèbre des opérateurs linéaires $\mathrm{Ca} \to \mathrm{Ca}$.

Etant donné que pour tout élément $u = \lambda + u' \in \mathrm{Ca} \subset \mathrm{Cl}(\mathrm{Ca}')$ on a $R(\lambda + u') = \lambda E + R_{u'} = R_{\lambda + u'}$, l'homomorphisme R est le prolongement de la correspondance $u \to R_u$ pour tout $u \in \mathrm{Ca}$.

On démontre de façon analogue l'existence de l'homomorphisme

$$L\colon \mathrm{Cl}(\mathrm{Ca}') \to \mathrm{End}\,\mathrm{Ca},$$

$L_u = L_u$: $X \to ux$ pour tout élément $u \in \mathrm{Ca}$.

Etant donné que l'algèbre $\mathrm{Cl}(\mathrm{Ca})$ s'identifie à l'algèbre $\mathrm{Cl}_+(8)$ et l'algèbre End Ca, à l'algèbre $\mathbb{R}(8)$, on obtient les deux homomorphismes composés

$$(4) \qquad \mathrm{Spin}_+(8) \subset \mathrm{Cl}_+^0(8) = \mathrm{Cl}_+^0(\mathrm{Ca}) \xrightarrow{\omega_+^{-1}} \mathrm{Cl}(\mathrm{Ca}') \overset{R,\,L}{\rightrightarrows} \mathrm{End}\ \mathrm{Ca} = \mathbb{R}(8).$$

Comme $\omega_+^{-1}(e_0 u) = u$, à tout élément du groupe Spin(8) de la forme $e_0 u$, où $u \in S^7 \subset \mathrm{Ca}$, est associé l'opérateur R_u: $\mathrm{Ca} \to \mathrm{Ca}$ par l'un de ces homomorphismes, et l'opérateur L_u: $\mathrm{Ca} \to \mathrm{Ca}$, par l'autre. Ces opérateurs sont orthogonaux, puisque $\mid u \mid = 1$ et l'algèbre Ca est normée. Les éléments $e_0 u$, $u \in S^7$, engendrant le groupe $\mathrm{Spin}_+(8)$, il s'ensuit que les homomorphismes (4) transforment le groupe $\mathrm{Spin}_+(8)$ en le groupe SO(8) (ou, si l'on n'effectue pas la dernière identification, en le groupe Ort Ca des opérateurs orthogonaux $\mathrm{Ca} \to \mathrm{Ca}$).

Les éléments du groupe SO(8) (ou Ort Ca) qui sont les images de $a \in \mathrm{Spin}_+(8)$ par les homomorphismes (4) seront désignés respectivement par a^R et a^L.

D'après les calculs effectués ci-dessus, si $a = (e_0 u_1) \ldots (e_0 u_r)$, où $u_1, \ldots, u_r \in S^7$, alors

$$a^R = R_{u_1} \circ \ldots \circ R_{u_r} \quad \text{et} \quad a^L = L_{u_1} \circ \ldots \circ L_{u_r}.$$

Remarque 2. De la théorie générale des représentations matricielles des groupes Spin(n), il s'ensuit immédiatement que les homomorphismes $a \mapsto a^R$ et $a \mapsto a^L$ sont équivalents aux représenta-

tions semi-spinorielles du groupe Spin(8) $\approx$ Spin$_+$(8) (cf. formule (20), leçon 13 pour $m = 1$), représentations qui, notons-le, sont des revêtements (visiblement à deux feuillets), car elles ont un noyau discret. Du reste, on peut prouver directement cette équivalence (il faut simplement se rappeler qu'on obtiendrait exactement les représentations (20) de la leçon 13 si, en identifiant $\mathrm{Cl}_*^0(8)$ à $\mathrm{Cl}_*^0(\mathrm{Ca})$, on prend pour base de l'algèbre Ca les éléments $1, e, h, g, i, f, -j, k$, et, en identifiant End Ca à $\mathbb{R}(8)$, les éléments $-1, i, -j, -k, e, f, g, h$). Mais comme nous n'aurons pas besoin de cette équivalence, nous glisserons sur sa démonstration.

Outre les homomorphismes $a \mapsto a^R$ et $a \mapsto a^L$, nous avons encore affaire à l'homomorphisme φ_0: Spin$_+$(8) $\to$ SO(8) de la proposition 4 de la leçon 13 qui à tout élément $a \in$ Spin$_+$(8) associe un opérateur orthogonal $\varphi_0(a): x \mapsto ax\bar{a}$ de Ca dans Ca. Nous désignerons cet opérateur par a^T. Comme Ker $\varphi_0 = \{1, -1\}$, on a $(-a)^T = a^T$, et, en particulier, $(e_0u)^T = (ue_0)^T$. D'autre part, ainsi qu'on l'a montré dans la démonstration de la proposition 4 de la leçon 13, $(ue_0)^T = u^\perp e_0^\perp$, où $u^\perp$ est la symétrie par rapport à l'hyperplan orthogonal au vecteur u. Mais, en vertu de la formule (2) de la leçon précédente, pour toutes octaves x, $u \in$ Ca on a $uxu = \bar{\bar{u}}xu =$ $= (2(u, \bar{x}) - \overline{\bar{x}u})u$ (la conjugaison est celle de Ca), d'où il s'ensuit pour $|u| = 1$ que $uxu = -u^\perp(\bar{x})$, c'est-à-dire — puisque $e_0^\perp x = -\bar{x}$ — que $uxu = u^\perp e_0^\perp(x)$. En considérant l'opérateur T_u: $x \mapsto uxu$, c'est-à-dire l'opérateur $T_u = R_u \circ L_u = L_u \circ R_u$, on trouve donc que $T_u = u^\perp e_0^\perp$ et, par suite, $T_u = (ue_0)^T = (e_0u)^T$. Par conséquent, si $a = (e_0u_1) \ldots (e_0u_r)$, alors

$$a^T = T_{u_1} \circ \ldots \circ T_{u_r}$$

comme pour les opérateurs a^R et a^L.

Lemme 2 (Identité centrale de R. Moufang). *Dans toute algèbre alternative on a l'identité*

$$u \cdot xy \cdot u = ux \cdot yu. \tag{5}$$

Démonstration. Les identités alternatives et d'élasticité entraînent

$$y^2x \cdot y = (y \cdot yx)y = y(yx \cdot y) = y(y \cdot xy) = y^2 \cdot xy$$

et, par suite, puisque les associateurs sont antisymétriques,

$$xy^2 \cdot y = x \cdot y^3,$$

c'est-à-dire que

$$x \cdot y^3 = (xy \cdot y)y.$$

En polarisant cette identité par rapport à y, c'est-à-dire en posant $y = a + b$ et en réduisant les termes semblables, on obtient l'iden-

tité

$$x\cdot a^2b + x\cdot ba^2 + x\cdot aba + x\cdot bab + x\cdot ab^2 + x\cdot b^2a =$$
$$= xa^2\cdot b + xb\cdot a^2 + (xa\cdot b)a + (xb\cdot a)b + xa\cdot b^2 + xb^2\cdot a.$$

Or les sommes des deux premiers termes de chaque membre sont égales à cause de l'antisymétrie des associateurs. *Idem* pour les sommes des deux derniers termes. Donc,

$$x\cdot aba + x\cdot bab = (xa\cdot b)a + (xb\cdot a)b.$$

En remplaçant b par λb, où $\lambda \in \mathbb{R}$, en simplifiant par λ et en faisant $\lambda = 0$, on obtient l'identité

$$x\cdot aba = (xa\cdot b)a,$$

appelée *identité droite de Moufang*.

En vertu de cette identité et de l'antisymétrie des associateurs

$$ab\cdot xa - a\cdot bx\cdot a = ab\cdot xa - (ab\cdot x)a + (ab\cdot x - a\cdot bx)a =$$
$$= (x\cdot ab)a - x\cdot aba + (xa\cdot b - x\cdot ab)a =$$
$$= (xa\cdot b)a - x\cdot aba = 0,$$

et l'on retrouve l'identité (5) aux notations près. □

Le lemme 2 exprime que

$$T_u(xy) = L_ux\cdot R_uy.$$

De là, on déduit par une récurrence évidente, en se servant des formules établies pour a^R, a^L et a^T, que

$$a^T(xy) = a^Lx\cdot a^Ry$$

pour toutes octaves $x, y \in \mathbb{C}\mathrm{a}$ et tout élément $a \in \mathrm{Spin}_+(8)$.

Nous pouvons maintenant passer du groupe $\mathrm{Spin}_+(8)$ au groupe Spin(8) qui lui est isomorphe. En désignant par a_+ l'élément du groupe $\mathrm{Spin}_+(8)$ associé à l'élément a du groupe Spin(8) et en posant $a^K = a_+^K$, où $K = T, L, R$, on obtient la proposition suivante :

Proposition 2. *Pour tout élément* $a \in \mathrm{Spin}(8)$, *on a l'identité*

$$a^T(xy) = a^Lx\cdot a^Ry, \quad x, y \in \mathbb{C}\mathrm{a}.$$

Cette proposition s'appelle *principe de trialité pour le groupe* Spin(8).

Remarque 3. On sait que l'homomorphisme $a \mapsto a^T$ est un revêtement, et en vertu de la remarque 1, qu'il en est de même des homomorphismes $a \mapsto a^L$ et $a \mapsto a^R$. Donc, pour tout $K = T, L, R$, l'homomorphisme $a \mapsto a^K$ induit un isomorphisme de l'algèbre de Lie $\mathfrak{l}(\mathrm{Spin}\,(8))$ sur l'algèbre de Lie $\mathfrak{l}(\mathrm{SO}(8)) = \mathfrak{so}(8)$. Si on

identifie ces algèbres de Lie à l'aide du premier de ces isomorphismes, les deux autres se transforment en automorphismes de l'algèbre de Lie $\mathfrak{so}(8)$. En désignant par A^λ et A^ρ les images d'une matrice quelconque $A \in \mathfrak{so}(8)$ par ces automorphismes, on aura de toute évidence l'identité

(6) $$A^\lambda x \cdot y + x \cdot A^\rho y = A(xy), \quad x, y \in \mathrm{Ca}, \quad A \in \mathfrak{so}(8).$$

Cette identité s'appelle *principe infinitésimal de trialité*. Sa démonstration est accessible dans l'article de Freudenthal [12].

La proposition suivante peut être regardée comme la réciproque du principe de trialité.

Proposition 3. *Si des opérateurs orthogonaux A, B, C: $\mathrm{Ca} \to \mathrm{Ca}$ sont tels que*

$$A(xy) = Bx \cdot Cy$$

pour toutes octaves $x, y \in \mathrm{Ca}$, alors ils sont nécessairement unimodulaires (appartiennent au groupe $\mathrm{SO}(8)$) et dans le groupe $\mathrm{Spin}(8)$ il existe un élément a et un seul tel que

(7) $$A = a^T, \quad B = a^L, \quad C = a^R.$$

Démonstration. Si $A \in \mathrm{SO}(8)$ (donc il existe un élément $a \in \mathrm{Spin}(8)$ tel que $A = a^T$), alors $xy = B'x \cdot C'y$, où $B' = (a^L)^{-1} \circ \circ B$ et $C' = (a^R)^{-1} \circ C$. Mais alors $B'x = xb$, où $b = (C'1)^{-1}$ et $C'y = cy$, où $c = (B'1)^{-1}$ et, par suite,

$$xy = xb \cdot cy.$$

En posant $x = y = 1$, on obtient que $bc = 1$, et en remplaçant x par $xc = xb^{-1}$, que $xc \cdot y = x \cdot cy$. Mais il est immédiat de voir (en prenant pour x et y tous les éléments de base de l'espace vectoriel Ca') que pour x et y quelconques, cette égalité n'est possible que si $c \in \mathbb{R}$, c'est-à-dire puisque C' est orthogonal, si $c = \pm 1$. Si $c = 1$, l'élément a vérifie les relations (7), si $c = -1$, il faut le remplacer par $-a$. L'unicité de cet élément est évidente (car si $a^T = b^T$, alors $a = \pm b$ et, par suite, $a^L = \pm b^L$).

A noter que les opérateurs B et C appartiennent au groupe $\mathrm{SO}(8)$. Donc, pour achever la démonstration il reste simplement à prouver que la relation d'appartenance $A \in \mathrm{O}(8) \setminus \mathrm{SO}(8)$ est impossible. Sans perdre en généralité, on peut de toute évidence admettre que A: $x \to \bar{x}$, c'est-à-dire que $Bx \cdot Cy = \overline{xy}$. Mais alors $Bx = \overline{xb}$, où $b = (C1)^{-1}$ et $Cy = \overline{cy}$, où $c = (B1)^{-1}$, c'est-à-dire que $\overline{xb} \cdot \overline{cy} = \overline{xy}$ et, par suite (on remplace x par $\bar{x}$ et y par $\bar{y}$),

$$xb \cdot cy = yx.$$

Pour $x = y = 1$, on en déduit que $bc = 1$ et, par suite, que $x \cdot cy = y \cdot xc$. En particulier, $cy = yc$ pour tous les $y \in \mathrm{Ca}$, ce qui n'est

possible que pour $c \in \mathbb{R}$. Donc, $xy = yx$, ce qui est absurde. Par conséquent, la relation d'appartenance $A \in \mathrm{O}(8) \setminus \mathrm{SO}(8)$ est impossible.

Remarque 4. Le principe infinitésimal de trialité (6) admet aussi une réciproque: *si A, B, C sont des opérateurs antisymétriques* $\mathbb{C}\mathrm{a} \to \mathbb{C}\mathrm{a}$ *tels que*

$$A(xy) = Bx \cdot y + x \cdot Cy \quad \forall x,\ y \in \mathbb{C}\mathrm{a},$$

alors $B = A^{\lambda}$ et $C = A^{\rho}$. En effet, les opérateurs $B' = B - A^{\lambda}$ et $C' = C - A^{\rho}$ vérifient l'identité $B'x \cdot y = -x \cdot C'y$, d'où il résulte que $B'x = xb$, où $b = -C'1$ et $C'y = cy$, où $y = -B'1$. Donc, $xb \cdot y = -x \cdot cy$, d'où l'on déduit pour $x = y = 1$ que $b = -c$, et donc, que $xb \cdot y = x \cdot by$. Par conséquent, $b \in \mathbb{R}$, c'est-à-dire que l'opérateur B' est diagonal. Donc $B' = 0$ par antisymétrie et, par suite, $C' = 0$. □

Le groupe Spin(9) est justiciable aussi d'un principe de trialité. Pour l'obtenir nous introduisons les espaces vectoriels $\mathbb{C}\mathrm{a}^{\oplus} = \mathbb{R} \oplus \mathbb{C}\mathrm{a}$ et $\mathbb{C}\mathrm{a}^2 = \mathbb{C}\mathrm{a} \oplus \mathbb{C}\mathrm{a}$. Pour éviter toute confusion avec le produit scalaire, on désignera les éléments de l'espace vectoriel $\mathbb{C}\mathrm{a}^2$ par $\{\xi, \eta\}$, où $\xi, \eta \in \mathbb{C}\mathrm{a}$, et respectivement, les éléments de l'espace vectoriel $\mathbb{C}\mathrm{a}^{\oplus}$, par les symboles $\{r, \rho\}$, où $r \in \mathbb{R}$, $\rho \in \mathbb{C}\mathrm{a}$.

A tout élément $x = \{\xi, \eta\}$ de $\mathbb{C}\mathrm{a}^2$ et tout élément $u = \{r, \rho\}$ de $\mathbb{C}\mathrm{a}^{\oplus}$ associons l'élément xu de $\mathbb{C}\mathrm{a}^2$ défini par la formule

$$xu = \{-r\xi + \bar{\rho}\bar{\eta},\ r\eta + \bar{\xi}\bar{\rho}\}.$$

Il est clair que cette multiplication est bilinéaire (sur $\mathbb{R}$) et possède la propriété suivante: si $xu = xv$ pour tout $x \in \mathbb{C}\mathrm{a}^2$, alors $u = v$ (si $-r\xi + \bar{\rho}\bar{\eta} = -s\xi + \bar{\sigma}\bar{\eta}$ pour tous ξ et η, alors $-r\xi = -s\xi$, donc $r = s$, et $\bar{\rho}\bar{\eta} = \bar{\sigma}\bar{\eta}$, donc $\rho = \sigma$).

Lemme 3. *Pour tous éléments $x \in \mathbb{C}\mathrm{a}^2$ et $u, v \in \mathbb{C}\mathrm{a}^{\oplus}$, on a l'identité*

$$xu \cdot v = -xv \cdot v^{\perp}u, \tag{8}$$

où comme toujours $v^{\perp}$ désigne la symétrie dans l'hyperplan perpendiculaire au vecteur v.

Démonstration. Soient $u = \{r, \rho\}$, $v = \{s, \sigma\}$. Il est clair que sans nuire à la généralité, on peut admettre que les vecteurs u et v sont des vecteurs unités, c'est-à-dire vérifiant les relations $r^2 + |\rho|^2 = 1$, $s^2 + |\sigma|^2 = 1$. Alors, $v^{\perp}u = \{r - 2(u, v)s,\ \rho - 2(u, v)\sigma\}$, où $2(u, v) = 2rs + \bar{\rho}\sigma + \bar{\sigma}\rho$.

A cause de la linéarité, il suffit de prouver l'identité (8) pour $x = \{\xi, 0\}$ et $x = \{0, \eta\}$. Mais si $x = \{\xi, 0\}$, alors

$$xu = \{-r\xi, \overline{\bar{\xi}\bar{\rho}}\}, \quad xv = \{-s\xi, \overline{\bar{\xi}\bar{\sigma}}\},$$
$$xu\cdot v = \{rs\xi + \bar{\sigma}\cdot\rho\xi, s\overline{\bar{\xi}\rho} - r\bar{\xi}\bar{\sigma}\}$$

et

$$xv\cdot v^{\perp}u = \{(r - 2(u, v)\,s)\,s\xi +$$
$$+ (\overline{\rho - 2(u, v)\,\sigma})\cdot\sigma\xi, (r - 2(u, v)\,s)\,\overline{\bar{\xi}\bar{\sigma}} - s\bar{\xi}(\overline{\rho - 2(u, v)\,\sigma})\} =$$
$$= \{rs\xi - 2(u, v)\,s^2\xi + \bar{\rho}\cdot\sigma\xi - 2(u, v)\,\bar{\sigma}\sigma\xi, r\overline{\bar{\xi}\,\sigma} -$$
$$- 2(u, v)\,\overline{s\bar{\xi}}\,\bar{\sigma} - s\overline{\bar{\xi}\,\rho} + 2(u, v)\,s\bar{\xi}\,\bar{\sigma}\} =$$
$$= \{[rs - 2(u, v)\,(s^2 + |\sigma|^2)]\,\xi + \bar{\rho}\cdot\sigma\xi, r\bar{\xi}\,\bar{\sigma} - s\overline{\bar{\xi}\rho}\},$$

et, par suite (puisque $rs - 2(u, v)\,(s^2 + |\sigma|^2) = rs - 2(u, v) = - rs - \bar{\rho}\sigma - \bar{\sigma}\rho$),

$$xu\cdot v + xv\cdot v^{\perp}u = \{\bar{\sigma}\cdot\rho\xi - \bar{\sigma}\rho\cdot\xi + \bar{\rho}\sigma\cdot\xi - \bar{\rho}\cdot\sigma\xi, 0\},$$

ce qui est égal à $\{0, 0\}$ en raison de l'antisymétrie des associateurs et de la remarque évidente qui dit que l'associateur change de signe quand on remplace un élément par son conjugué.

Le cas $x = \{0, \eta\}$ se traite de façon analogue. □

Associons maintenant à tout élément $u \in \mathrm{Ca}^{\oplus}$ l'opérateur linéaire $R_u\colon x \mapsto xu$ de Ca^2 dans Ca^2. Comme, ainsi que le montre un calcul immédiat, $xu\cdot u = |u^2|\,x$, $\forall x \in \mathrm{Ca}^2$, $\forall u \in \mathrm{Ca}^{\oplus}$, c'est-à-dire que $R_u^2 = |u^2|\,E$, il vient que l'application linéaire $u \mapsto R_u$ de l'espace $\mathrm{Ca}^{\oplus}$ dans l'algèbre des opérateurs linéaires End Ca^2 se prolonge de façon unique en un homomorphisme d'algèbres

$$R\colon \mathrm{Cl}_+(\mathrm{Ca}^{\oplus}) \to \mathrm{End}\,\mathrm{Ca}^2.$$

L'identification des algèbres $\mathrm{Cl}_+(\mathrm{Ca}^{\oplus})$ et End Ca^2 respectivement aux algèbres $\mathrm{Cl}_+(9)$ et $\mathbb{R}(16)$ fait apparaître l'homomorphisme composé

$$\text{(9)} \qquad \mathrm{Spin}_+(9) \subset \mathrm{Cl}_+(9) = \mathrm{Cl}(\mathrm{Ca}^{\oplus}) \xrightarrow{R} \mathrm{End}\,\mathrm{Ca}^2 = \mathbb{R}(16).$$

Si $u = \{r, \rho\} \in S^8 \subset \mathrm{Ca}^{\oplus}$, i.e. $r^2 + |\rho|^2 = 1$ et $x = \{\xi, \eta\} \in \mathrm{Ca}^2$, alors

$$|xu|^2 = |\{-r\xi + \bar{\rho}\bar{\eta}, r\eta + \overline{\bar{\xi}\bar{\rho}}\}|^2 =$$
$$= (-r\xi + \bar{\rho}\bar{\eta})\,(-r\bar{\xi} + \eta\rho) + (r\eta + \overline{\bar{\xi}\bar{\rho}})\,(r\bar{\eta} + \rho\xi) =$$
$$= r^2\xi\bar{\xi} - r\xi\cdot\eta\rho - \overline{r\rho\eta}\cdot\bar{\xi} + \bar{\rho}\,\bar{\eta}\cdot\eta\rho +$$
$$+ r^2\eta\bar{\eta} + r\eta\cdot\rho\xi + r\bar{\xi}\,\bar{\rho}\cdot\bar{\eta} + \overline{\bar{\xi}\bar{\rho}}\cdot\rho\xi =$$
$$= (r^2 + |\rho|^2)\,(|\xi|^2 + |\eta|^2) + rz = |x|^2 + rz,$$

où

$$z = -\xi\cdot\eta\rho - \overline{\rho\eta}\cdot\overline{\xi} + \eta\cdot\rho\xi + \overline{\xi}\,\overline{\rho}\cdot\overline{\eta}.$$

Mais en posant $[\xi, \eta] = \xi\eta - \eta\xi$ et $(\xi, \eta, \zeta) = \xi\eta\cdot\zeta - \xi\cdot\eta\zeta$, on déduit sans peine que

$$z = [\eta\rho, \xi] - (\eta, \rho, \xi) + [\overline{\xi}, \overline{\rho\eta}] + (\overline{\xi}, \overline{\rho}, \overline{\eta}).$$

Comme déjà signalé, la substitution à tout élément de son conjugué change le signe de l'associateur (ξ, η, ζ). Donc, $(\overline{\xi}, \overline{\rho}, \overline{\eta}) = -(\xi, \rho, \eta)$. De façon analogue, $[\overline{\xi}, \overline{\rho\eta}] = [\overline{\xi}, \overline{\eta}\,\overline{\rho}] = [\xi, \eta\rho]$. Donc,

$$z = [\eta\rho, \xi] + [\xi, \eta\rho] - (\eta, \rho, \xi) - (\xi, \rho, \eta) = 0$$

à cause de l'antisymétrie des associateurs. Donc, $|xu|^2 = |x|^2$. Ceci exprime que l'opérateur linéaire $R_u = Ru$ est orthogonal, d'où il résulte immédiatement que l'homomorphisme (9) envoie le groupe $\mathrm{Spin}_+(9)$ (et même le groupe $\mathrm{pin}_+(9)$) dans le groupe SO(9).

L'image d'un élément $a \in \mathrm{Spin}_+(9)$ dans le groupe SO(9) par l'homomorphisme (9), traitée comme un opérateur orthogonal de Ca^2 dans Ca^2, sera désignée par a^R.

Remarque 5. L'homomorphisme $a \to a^R$ est une représentation spinorielle du groupe $\mathrm{Spin}_+(9) \approx \mathrm{Spin}(9)$, mais nous ne prouverons pas ce fait, car nous n'en aurons pas besoin.

L'image d'un élément $a \in \mathrm{Spin}_+(9)$ par l'homomorphisme φ_0: $\mathrm{Spin}_+(9) \to \mathrm{SO}(9)$ de la proposition 4 de la leçon 13, traitée comme un opérateur orthogonal de $\mathrm{Ca}^{\oplus}$ dans $\mathrm{Ca}^{\oplus}$, sera désignée par a^T.

Par ailleurs, comme plus haut, pour tout élément $a \in \mathrm{Spin}(9)$, posons $a^R = a_+^R$ et $a^T = a_+^T$, où a_+ est l'élément du groupe $\mathrm{Spin}_+(9)$ associé à l'élément a par l'isomorphisme $\mathrm{Spin}(9) \approx \mathrm{Spin}_+(9)$. On aura alors la proposition suivante:

Proposition 4. *Pour tout élément* $a \in \mathrm{Spin}(9)$, *on a l'identité*

(10) $$a^R(xu) = a^R x\cdot a^T u, \quad x \in \mathrm{Ca}^2, \quad u \in \mathrm{Ca}^{\oplus}.$$

D é m o n s t r a t i o n. Le groupe Spin(9) étant engendré par les éléments vw, où $v, w \in S^8 \subset \mathrm{Ca}^{\oplus}$, il suffit de prouver l'identité (10) pour $a = vw$. Or, en vertu du lemme 2 et de la définition de l'homomorphisme R

$$\begin{aligned}(vw)^R(xu) &= (v^R \circ w^R)(xu) = (R_v \circ R_w)(xu) = \\ &= R_v(xu\cdot w) = R_v(-xw\cdot w^{\perp}u) = \\ &= -(xw\cdot w^{\perp}u)\cdot v = (xw\cdot v)\cdot v^{\perp}(w^{\perp}u) = \\ &= (R_v \circ R_w)\,x\cdot(v^{\perp} \circ w^{\perp})\,u = \\ &= (vw)^R x\cdot(vw)^T u\end{aligned}$$

(on rappelle que $(vw)^T = v^{\perp} \circ w^{\perp}$). □

Nous avons la proposition suivante, identique à la proposition 3:

Proposition 5. *Si des opérateurs orthogonaux* A: $\mathrm{Ca}^{\oplus} \to \mathrm{Ca}^{\oplus}$ *et* B: $\mathrm{Ca}^2 \to \mathrm{Ca}^2$ *sont tels que*

$$B(xu) = Bx \cdot Au \tag{11}$$

pour tous éléments $x \in \mathrm{Ca}^2$ *et* $u \in \mathrm{Ca}^{\oplus}$, *alors ces opérateurs sont nécessairement unimodulaires et il existe un élément* $a \in \mathrm{Spin}(9)$ *et un seul tel que*

$$A = a^T, \ B = a^R. \tag{12}$$

Prouvons préalablement le lemme suivant:

Lemme 4. *Si* B *est un opérateur orthogonal* $\mathrm{Ca}^2 \to \mathrm{Ca}^2$, *tel que*

$$B(xu) = Bx \cdot u \tag{13}$$

pour tous éléments $x \in \mathrm{Ca}^2$ *et* $u \in \mathrm{Ca}$ (*c'est-à-dire pour* $u = \{0, \rho\}$, *où* $\rho \in \mathrm{Ca}$), *alors* $B = \pm E$.

Démonstration. Supposons que

$$Bx = \begin{cases} \{C\xi, C_1\xi\} & \text{pour} \quad x = \{\xi, 0\}, \\ \{D_1\eta, D\eta\} & \text{pour} \quad x = \{0, \eta\} \end{cases}$$

et, par suite,

$$Bx = \{C\xi + D_1\eta, C_1\xi + D\eta\} \quad \text{pour} \quad x \in \{\xi, \eta\}.$$

Dans ces notations, la relation (13) devient pour $u = \{0, \rho\}$

$$\{C(\overline{\rho\eta}) + D_1(\bar{\xi}\bar{\rho}), C_1(\bar{\rho}\eta) + D(\bar{\xi}, \bar{\rho})\} = \{\bar{\rho}(\overline{C_1\xi + D\eta}), (\overline{C\xi + D_1\eta})\bar{\rho}\},$$

d'où il s'ensuit (on admet d'abord que $\xi = 0$ et ensuite que $\eta = 0$) que

$$C(\overline{\bar{\rho}\eta}) = \bar{\rho} \cdot \overline{D\eta}, \qquad C_1(\bar{\rho}\bar{\eta}) = \overline{D_1\eta} \cdot \bar{\rho},$$

$$D_1(\bar{\xi}\bar{\rho}) = \bar{\rho} \cdot \overline{C_1\xi}, \qquad D(\bar{\xi}\bar{\rho}) = \overline{C\xi} \cdot \bar{\rho}.$$

Pour $\rho = 1$, on en déduit que

$$D_1\bar{\xi} = \overline{C_1\xi}, \qquad D\bar{\xi} = \overline{C\xi}$$

et par conséquent (on remplace $\bar{\rho}$ par ρ et $\bar{\xi}$ par ξ) que

$$C(\rho\xi) = \rho \cdot C\xi, \quad C_1(\rho\xi) = C_1\xi \cdot \rho. \tag{14}$$

Donc $C\rho = \rho c$ et $C_1\rho = c_1\rho$, où $c = C1$, $c_1 = C_1 1$. Les identités (14) deviennent alors

$$\rho\xi \cdot c = \rho \cdot \xi c, \quad c_1 \cdot \rho\xi = c_1\rho \cdot \xi.$$

On sait déjà que la première identité entraîne que $c \in \mathbb{R}$. En posant $c_1 = u + ve$, $\rho = w$ et $\xi = e$, où $u, v, w \in \mathbb{H}$, dans la se-

conde, et en tenant compte de ce que $(u + ve)\cdot we = -\overline{w}v + wu\cdot e$ et $(u + ve)e\cdot w = -\overline{w}u - wv\cdot e$, on obtient les relations $\overline{w}v = \overline{u}u$, $wu = -wv$, qui, w étant quelconque, ne sont possibles que pour $u = v = 0$, c'est-à-dire pour $c_1 = 0$. Ceci prouve que $C_1 = 0$ (donc que $D_1 = 0$) et que l'opérateur C est confondu avec l'opérateur D et est un opérateur de multiplication par le réel c. Donc, l'opérateur B est aussi un opérateur de multiplication par c, ce qui n'est possible que pour $c = \pm 1$. □

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 5. Si l'opérateur A est unimodulaire et, donc, il existe un élément $a \in \in \mathrm{Spin}(9)$ tel que $A = a^T$, alors l'opérateur $B \circ (a^R)^{-1}$ satisfait la relation (13) (même pour u quelconque) et, par suite, $B = \pm a^R = (\pm a)^R$ en vertu du lemme 4. Comme $(-a)^T = a^T$, ceci prouve l'égalité (11). L'unicité de l'élément a dans ces égalités résulte du fait que l'égalité $b^T = a^T$ n'est possible que pour $b = \pm a$ et $(-a)^R = -a^R \neq a^R$.

Donc, pour achever la démonstration de la proposition 5, il nous reste à prouver que l'identité (11) n'a lieu que pour un opérateur A unimodulaire. Mais si cette identité est satisfaite pour un opérateur orthogonal A non unimodulaire, elle le sera visiblement pour tout autre opérateur tel que A (avec un autre B bien sûr). Il suffit donc de montrer l'absurdité de la proposition suivante: l'identité (11) a lieu pour l'opérateur A: $u \mapsto \bar{u}$, où $\bar{u} = \{-r, \rho\}$ si $u = \{r, \rho\}$, c'est-à-dire qu'existe un opérateur orthogonal B: $\mathrm{Ca}^2 \to \mathrm{Ca}^2$ tel que

$$B(xu) = Bx\cdot\bar{u}. \tag{15}$$

Or, si $u = \{0, \rho\}$ l'identité (15) est confondue avec l'identité (13), donc $B = \pm E$ en vertu du lemme 4. Donc, l'identité (15) équivaut à l'identité $xu = x\bar{u}$, d'où il s'ensuit que $u = \bar{u}$, ce qui est absurde. Ceci achève la démonstration de la proposition 5. □

Outre les octaves, on peut par exemple étudier des matrice dont les éléments sont des octaves. Etant donné que l'algèbre Ca est munie de la conjugaison, toute matrice d'octaves X possède une *matrice transposée conjuguée* X^*. Par analogie avec le cas complexe une matrice d'octaves X est *hermitienne* si $X^* = X$.

Définissant le produit XY des matrices d'octaves X et Y par la formule usuelle, on obtient immédiatement (par des calculs identiques à ceux effectués pour les matrices complexes) que pour toutes matrices d'octaves X et Y, on a, comme pour les matrices complexes, l'égalité $(XY)^* = Y^*X^*$ qui entraîne que les matrices hermitiennes d'octaves d'ordre n forment une algèbre pour la *multiplication jordanienne*

$$X \circ Y = \frac{XY + YX}{2}.$$

Cette algèbre est commutative et unitaire. Son unité est la matrice unité E.

Etudions cette algèbre pour $n = 3$. Désignons-la par Al en l'honneur du mathématicien américain A. Albert qui fut parmi les premiers à la remarquer.

Remarque 6. La multiplication jordanienne a un sens sur toute algèbre. Elle est visiblement commutative, et, sur une algèbre associative, satisfait l'identité

$$(x^2 \circ y) \circ x = x^2 \circ (y \circ x), \tag{16}$$

où $x^2 = x \circ x = xx$. Les algèbres à multiplication commutative vérifiant cette identité s'appellent *algèbres de Jordan*. (A propos, nous avons vu en démontrant le lemme 1 que l'identité (16) est réalisée sur toute algèbre alternative; donc, une algèbre alternative commutative est une algèbre de Jordan.) De toute évidence, puisque l'algèbre des octaves n'est pas associative, il n'y a aucune raison pour que l'algèbre des matrices d'octaves hermitiennes d'ordre n soit algèbre de Jordan. Néanmoins, il s'avère que c'est parce que l'algèbre Ca est alternative que cette algèbre est de Jordan pour $n \leqslant 3$, et en outre, comme l'a montré Albert pour $n = 3$, elle ne peut être déduite d'aucune algèbre associative. Ceci explique le rôle particulier de l'algèbre Al et l'intérêt que nous lui portons. Cependant nous glisserons sur la démonstration du fait que cette matrice est de Jordan, car nous n'en aurons pas besoin.

Tout élément X de l'algèbre Al se représente d'une seule façon sous la forme

$$X = a_1E_1 + a_2E_2 + a_3E_3 + X_1(\xi_1) + X_2(\xi_2) + X_3(\xi_3), \tag{17}$$

où

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

et

$$X_1(\xi) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \xi \\ 0 & \bar\xi & 0 \end{pmatrix}, \quad X_2(\xi) = \begin{pmatrix} 0 & 0 & \bar\xi \\ 0 & 0 & 0 \\ \xi & 0 & 0 \end{pmatrix}, \quad X_3(\xi) = \begin{pmatrix} 0 & \xi & 0 \\ \bar\xi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

et en outre, comme le montre un calcul immédiat,

$$E_i \circ E_j = \begin{cases} E_j & \text{pour } j = i, \\ 0 & \text{pour } j \neq i, \end{cases} \tag{18}$$

$$E_i \circ X_j(\xi) = \begin{cases} 0, & \text{pour } j = i, \\ \frac{1}{2} X_j(\xi) & \text{pour } j \neq i, \end{cases} \tag{19}$$

$$(20)\qquad X_i(\xi)\circ X_j(\eta)=\begin{cases}(\xi,\eta)(E-E_i) & \text{pour } j=i,\\ X_{i+2}(\overline{\xi\eta}) & \text{pour } j=i+1\end{cases}$$

(dans la dernière formule les indices sont considérés comme des résidus mod 3; comme, en vertu de cette convention, $i = j + 1$ pour $j = i + 2$, le cas $j = i + 2$ se ramène au cas $j = i + 1$).

Ceci définit entièrement la structure algébrique de l'algèbre Al.

L'une des plus importantes caractéristiques d'une algèbre est la structure de l'ensemble de ses *idempotents*, c'est-à-dire de ses éléments x tels que $x^2 = x$. Pour l'algèbre Al nous étudions l'ensemble de ses idempotents X tels que Tr $X = 1$. Ces idempotents seront appelés *idempotents primitifs*.

Pour un élément (17), la condition Tr $X = 1$ exprime que

$$(21)\qquad a_1 + a_2 + a_3 = 1,$$

et la condition $X^2 = X$ se ramène aux six équations suivantes:

$$\begin{aligned}&a_i^2 + |\xi_{i+1}|^2 + |\xi_{i+2}|^2 = a_i,\\ &\bar{\xi}_{i+2}\bar{\xi}_{i+1} + (a_{i+1} + a_{i+2})\,\xi_i = \xi_i,\end{aligned}\qquad i = 1,\ 2,\ 3 \bmod 3,$$

qui, en vertu de la condition (21), sont équivalentes aux équations

$$(22)\qquad a_i(a_{i+1} + a_{i+2}) = |\xi_{i+1}|^2 + |\xi_{i+2}|^2,$$

$$i = 1,\ 2,\ 3 \bmod 3.$$

$$(23)\qquad a_i\bar{\xi}_i = \xi_{i+1}\xi_{i+2},$$

Des équations (23) on déduit simultanément que $a_i|\xi_i|^2 = \xi_i\xi_{i+1}\xi_{i+2}$ et $a_i|\xi_i|^2 = \xi_{i+1}\xi_{i+2}\xi_i$. Par conséquent, le nombre réel $\lambda = \xi_i\xi_{i+1}\xi_{i+2}$ ne change pas si l'on remplace i par $i + 1$ et, par suite, est le même pour tous les i. Ceci étant,

$$(24)\qquad a_i|\xi_i|^2 = \lambda \quad \text{pour tout} \quad i = 1,\ 2,\ 3.$$

Maintenant il est immédiat de voir que

$$(25)\qquad |\xi_i|^2 = a_{i+1}a_{i+2} \quad \text{pour tout} \quad i = 1,\ 2,\ 3 \bmod 3.$$

En effet, en multipliant l'équation (22) par $a_{i+1}a_{i+2}$, on obtient en vertu de (24)

$$a_ia_{i+1}a_{i+2}(a_{i+1} + a_{i+2}) = a_{i+2}\lambda + a_{i+1}\lambda,$$

d'où l'on déduit pour $a_{i+1} + a_{i+2} \neq 0$, c'est-à-dire pour $a_i \neq 1$, que $\lambda = a_ia_{i+1}a_{i+2}$. Par conséquent, si par ailleurs $a_i \neq 0$, alors (25) résulte de (24). Si $a_i = 0$ ou $a_{i+1} + a_{i+2} = 0$, alors il s'ensuit immédiatement de (22) que $\xi_{i+1} = \xi_{i+2} = 0$, et donc (25) découle de (22) pour $i + 1$. □

Réciproquement, si les conditions (25) sont remplies pour tous les i, alors il en est de même des conditions (22). Ceci prouve qu'*une matrice* $X \in$ Al *est un idempotent primitif si et seulement si ses éléments satisfont les conditions* (21), (23) *et* (25).

On remarquera que de (25) et (21), il résulte immédiatement que $a_i > 0$ quel que soit $i = 1, 2, 3$.

On démontre sans peine que pour $\xi_1, \xi_2, \xi_3 \in \mathbb{C}$, les équations (21), (23) et (25) définissent une variété difféomorphe au plan projectif complexe $\mathbb{C}P^2$ (le difféomorphisme $[z_1 : z_2 : z_3] \mapsto (\xi_1, \xi_2, \xi_3, a_1, a_2, a_3)$ est défini par les formules

$$\xi_i = \frac{z_{i+1}\bar{z}_{i+2}}{|z_1|^2 + |z_2|^2 + |z_3|^2}, \quad a_i = \frac{|z_i|^2}{|z_1|^2 + |z_2|^2 + |z_3|^2},$$

où $i = 1, 2, 3 \bmod 3$), et pour $\xi_1, \xi_2, \xi_3 \in \mathbb{H}$, une variété difféomorphe au plan projectif quaternionique $\mathbb{H}P^2$ (l'espace projectif $\mathscr{A}P^n$ à n dimensions sur une algèbre associative $\mathscr{A}$ à division — et, en particulier, sur l'algèbre des quaternions $\mathbb{H}$ — se définit comme l'ensemble quotient de l'ensemble $\mathscr{A}^{n+1} \setminus \{0\}$ par la proportionnalité des vecteurs; l'associativité est nécessaire pour que cette relation soit transitive). Pour cette raison l'ensemble des idempotents primitifs de l'algèbre Al s'appelle *plan projectif des octaves* et est désigné par $\mathbb{C}$a P^2. (Il y a cependant des causes plus profondes à cette terminologie; on peut par exemple (cf. [12]) définir dans $\mathbb{C}$a P^2 des « droites » qui sont, en fait, des sphères à huit dimensions qui sont justiciables des axiomes d'incidence de géométrie projective; mais ceci déborde le cadre de notre exposé.)

Soit U_i, $i = 1, 2, 3$, un sous-ensemble ouvert du plan projectif $\mathbb{C}$a P^2, composé des points $(\xi_1, \xi_2, \xi_3, a_1, a_2, a_3)$ tels que $a_3 \neq 0$. Il est aisé de voir que cet ensemble est difféomorphe au produit $\mathbb{C}a \times \mathbb{C}a \approx \mathbb{R}^{16}$ (par exemple pour $i = 3$ le difféomorphisme $\mathbb{C}a \times \mathbb{C}a \to U_3$ est défini par les formules

$$\xi_1 = \frac{\eta_2}{1 + |\eta_1|^2 + |\eta_2|^2}, \quad \xi_2 = \frac{\eta_1}{1 + |\eta_1|^2 + |\eta_2|^2},$$

$$\xi_3 = \frac{\bar{\eta}_1\bar{\eta}_2}{1 + |\eta_1|^2 + |\eta_2|^2}, \quad a_1 = \frac{|\eta_1|^2}{1 + |\eta_1|^2 + |\eta_2|^2},$$

$$a_2 = \frac{|\eta_2|^2}{1 + |\eta_1|^2 + |\eta_2|^2}, \quad a_3 = \frac{1}{1 + |\eta_1|^2 + |\eta_2|^2},$$

où $\eta_1, \eta_2 \in \mathbb{C}a$) et, par suite, est simplement connexe. Comme les intersections $U_1 \cap U_2$ et $(U_1 \cap U_2) \cap U_3$ sont visiblement connexes (la première est difféomorphe au produit $\mathbb{C}a \times (\mathbb{C}a \setminus \{0\})$, la seconde, au produit $(\mathbb{C}a \setminus \{0\}) \times (\mathbb{C}a \setminus \{0\})$), $\mathbb{C}a\,P^2 = U_1 \cup U_2 \cup U_3$, d'où il découle que *le plan projectif des octaves* $\mathbb{C}$a P^2 *est une variété connexe et simplement connexe* (de dimension 16).

LEÇON 16

Produits scalaires sur l'algèbre Al.— Automorphismes et dérivations de l'algèbre Al.— Dérivations adjointes de l'algèbre Al.— Théorème de Freudenthal.— Corollaires du théorème de Freudenthal.— Groupe de Lie F_4.— Algèbre de Lie $\mathfrak{f}_4$.— Structure de l'algèbre de Lie $\mathfrak{f}_4^{\mathbb{C}}$.

Dire que tout élément X de l'algèbre Al satisfait l'égalité

$$X = a_1E_1 + a_2E_2 + a_3E_3 + X_1(\xi_1) + X_2(\xi_2) + X_3(\xi_3)$$

(cf. formule (17) de la leçon 15) revient à dire que l'espace vectoriel Al est somme directe de trois espaces vectoriels $\mathbb{R}$ et de trois espaces vectoriels Ca (de sorte que, en particulier, dim $Al = 3 + 3\cdot 8 = 27$). L'espace vectoriel Al est muni d'une structure euclidienne (d'un produit scalaire), car $\mathbb{R}$ et Ca sont euclidiens. De plus, la norme (la longueur) $|X|$ de tout élément X de cet espace vérifie la relation

$$(1) \qquad |X|^2 = a_1^2 + a_2^2 + a_3^2 + |\xi_1|^2 + |\xi_2|^2 + |\xi_3|^2.$$

Le produit scalaire (X, Y) des éléments X et Y de Al s'exprime en fonction de leur produit jordanien par la formule

$$(2) \qquad (X, Y) = \mathrm{Tr}(X \circ Y),$$

qui se vérifie par un calcul direct.

De cette formule, il résulte, en particulier, qu'*un idempotent* $X \in Al$ *est primitif si et seulement si* $|X| = 1$.

A noter que cette algèbre n'est pas métrique, car elle n'est munie d'aucune conjugaison. D'après le théorème de Hurwitz, elle ne peut être normée.

Outre le produit scalaire, l'algèbre Al est munie d'une importante fonctionnelle dont nous commencerons de loin la construction.

Etant donné qu'à cause de l'antisymétrie l'associateur $(ax)y -$

— $a(xy)$ change de signe par la permutation $(a, x, y) \mapsto (y, a, x)$, dans toute algèbre alternative on a l'identité

$$(ax)y + y(ax) = a(xy) + (ya)x.$$

En désignant a par a_i^j, x par x_j^k et y par y_k^i, on peut mettre cette identité sous la forme suivante:

$$(a_i^j x_j^k)\, y_k^i + y_k^i\, (a_i^j x_j^k) = a_i^j\, (x_j^k y_k^i) + (y_k^i a_i^j)\, x_j^k.$$

Cette identité reste valable si la sommation est réalisée sur tous les indices entre 1 et n (en fait on se limitera à $n = 3$). Mais en introduisant les matrices $A = (a_i^j)$, $X = (x_j^k)$ et $Y = (y_k^i)$, on peut mettre l'identité obtenue sous la forme

$$\text{(3)} \qquad \operatorname{Tr}(AX \cdot Y) + \operatorname{Tr}(Y \cdot AX) = \operatorname{Tr}(A \cdot XY) + \operatorname{Tr}(YA \cdot X),$$

qui est valable pour toutes matrices dont les éléments appartiennent à une algèbre alternative et donc, en particulier, pour les matrices de l'algèbre Al.

A noter que ce procédé est assez général et s'applique à toute identité multilinéaire. Par exemple, désignant par Re ξ le coefficient de 1 dans l'octave ξ, c'est-à-dire le produit scalaire $(\xi, 1)$, on aura l'identité

$$\operatorname{Re}(ab) = \operatorname{Re}(ba),$$

qui est valable pour toutes octaves a et b. Donc, les matrices d'octaves satisfont l'identité

$$\text{(4)} \qquad \operatorname{Re}\operatorname{Tr}(AB) = \operatorname{Re}\operatorname{Tr}(BA),$$

qui exprime que l'application de Re Tr rend commutative la multiplication des matrices d'octaves.

En vertu de cette commutativité, l'identité (3) entraîne

$$2\operatorname{Re}\operatorname{Tr}(AX \cdot Y) = \operatorname{Re}\operatorname{Tr}(A \cdot XY) + \operatorname{Re}\operatorname{Tr}(YA \cdot X).$$

Une permutation cyclique de A, X et Y nous donne l'identité

$$2\operatorname{Re}\operatorname{Tr}(XY \cdot A) = \operatorname{Re}\operatorname{Tr}(X \cdot YA) + \operatorname{Re}\operatorname{Tr}(AX \cdot Y).$$

En retranchant ces identités membre à membre et en se servant encore de la commutativité, on obtient après simplification par 3 l'identité

$$\text{(5)} \qquad \operatorname{Re}\operatorname{Tr}(AX \cdot Y) = \operatorname{Re}\operatorname{Tr}(A \cdot XY),$$

qui exprime que l'application de Re Tr rend associative la multiplication des matrices d'octaves.

Idem pour la multiplication jordanienne des matrices d'octaves:

$$\operatorname{Re}\operatorname{Tr}((A \circ X) \circ Y) = \operatorname{Re}\operatorname{Tr}((A \circ X) \circ Y) =$$

$$= \frac{1}{2}[\operatorname{Re}\operatorname{Tr}(AX \cdot Y) + \operatorname{Re}\operatorname{Tr}(Y \cdot XA)] =$$

$$= \frac{1}{2}[\mathrm{Re}\,\mathrm{Tr}(A \cdot XY) + \mathrm{Re}\,\mathrm{Tr}(YX \cdot A)] =$$
$$= \mathrm{Re}\,\mathrm{Tr}(A \cdot (X \circ Y)) = \mathrm{Re}\,\mathrm{Tr}(A \circ (X \circ Y)).$$

Comme la trace des matrices de Al est réelle, on en déduit en définitive que

$$\mathrm{Tr}((A \circ X) \circ Y) = \mathrm{Tr}(A \circ (X \circ Y))$$

pour toutes matrices A, X, $Y \in$ Al. D'après la formule (2), cela signifie que dans l'algèbre Al on a l'identité

$$(A \circ X, Y) = (A, X \circ Y),$$

qui entraîne (puisque l'algèbre Al est commutative) que *la fonctionnelle trilinéaire*

$$(6) \qquad (X, Y, Z) = \mathrm{Tr}((X \circ Y) \circ Z) = (X \circ Y, Z), \qquad X, Y, Z \in \mathrm{Al},$$

est symétrique en X, Y *et* Z.

Nous appellerons cette fonctionnelle *triproduit scalaire.*

Si un automorphisme Φ: Al → Al de l'algèbre Al préserve la trace, c'est-à-dire que $\mathrm{Tr}(\Phi X) = \mathrm{Tr}\, X$ pour tout élément $X \in$ Al, alors il préservera de toute évidence les deux produits scalaires

$$(7) \qquad (\Phi X, \Phi Y) = (X, Y) \quad \text{et} \quad (\Phi X, \Phi Y, \Phi Z) = (X, Y, Z)$$

quels que soient les éléments X, Y, $Z \in$ Al.

De là il résulte, en particulier, que le groupe des automorphismes de l'algèbre Al préservant la trace est un sous-groupe fermé du groupe orthogonal O(27) et donc est un groupe de Lie compact. Nous désignerons ce groupe par F_4 et son algèbre de Lie par $\mathfrak{f}_4$. Les éléments de l'algèbre de Lie $\mathfrak{f}_4$ sont les dérivations D: Al → Al qui annulent la trace, c'est-à-dire telles que $\mathrm{Tr}(DA) = 0$ pour toute matrice $A \in$ Al.

Remarque 1. On montrera ultérieurement que tout automorphisme préserve la trace, de sorte qu'en réalité le groupe F_4 est le groupe des automorphismes de l'algèbre Al, mais pour des raisons de commodité nous ajournons la démonstration de ce fait.

Il est remarquable que, réciproquement, *tout opérateur linéaire* Φ: Al → Al *respectant les deux produits scalaires est un automorphisme de l'algèbre* Al *préservant la trace.* En effet,

$$(\Phi X \circ \Phi Y - \Phi(X \circ Y), \Phi Z) =$$
$$= (\Phi X, \Phi Y, \Phi Z) - (\Phi(X \circ Y), \Phi Z) =$$
$$= (X, Y, Z) - (X \circ Y, Z) = 0$$

quels que soient X, Y, $Z \in$ Al, et donc $\Phi X \circ \Phi Y = \Phi(X \circ Y)$, car étant une isométrie pour le produit scalaire (2), l'opérateur Φ n'est

pas dégénéré et, par suite, tout élément de l'algèbre Al peut être mis sous la forme ΦZ. Ceci prouve que l'opérateur Φ est un automorphisme de l'algèbre Al. Donc $\Phi E = E$ et, par conséquent, $\mathrm{Tr}(\Phi X) = \mathrm{Tr}\, X$, car $\mathrm{Tr}\, X = (X, E)$ pour tout $X \in \mathrm{Al}$. □

Donc, indépendamment de l'algèbre Al, le groupe F_4 peut être traité comme le groupe des isométries d'un espace euclidien à 27 dimensions préservant une fonctionnelle trilinéaire.

On sait que pour les opérateurs Φ de la forme e^{tD}, la préservation du produit scalaire (de l'isométrie) équivaut à l'antisymétrie de l'opérateur D, c'est-à-dire qu'on a l'identité

$$(DX, Y) + (X, DY) = 0.$$

De façon analogue, la préservation du triproduit scalaire équivaut à la réalisation de l'identité

$$(DX, Y, Z) + (X, DY, Z) + (X, Y, DZ) = 0 \tag{8}$$

quels que soient les éléments $X, Y, Z \in \mathrm{Al}$. En effet, en dérivant la fonction $f(t) = (e^{tD}X, e^{tD}Y, e^{tD}Z)$ par rapport à t, on trouve immédiatement en raison de la linéarité que $f'(0)$ est égale au premier membre de l'identité (8). Donc, si $f(t) = \mathrm{const}$, alors (7) est réalisée. Réciproquement, si (8) est remplie, alors $f'(t) = 0$ pour tous les t, et donc $f(t) = \mathrm{const}$, ce qui équivaut à la deuxième identité (7). □

Pour des raisons évidentes, les opérateurs linéaires D vérifiant l'identité (8) seront appelés *antisymétriques pour le triproduit scalaire* (6).

D'après tout ce qui précède, *un opérateur linéaire* $D: \mathrm{Al} \to \mathrm{Al}$ *est une dérivation de l'algèbre* Al *annulant la trace* (*appartient à l'algèbre de Lie* $\mathfrak{f}_4$) *si et seulement s'il est antisymétrique pour les deux produits scalaires* (2) *et* (6).

Une matrice d'octaves A est *antihermitienne* si $A^* = -A$. Un calcul direct montre que si une matrice A est antihermitienne, alors pour toute matrice X hermitienne, la matrice $[A, X] = AX - XA$ est aussi hermitienne. Donc, toute matrice antihermitienne A d'ordre trois définit grâce à la formule

$$(\mathrm{ad}\, A)X = [A, X], \qquad X \in \mathrm{Al},$$

un opérateur linéaire $\mathrm{ad}\, A: \mathrm{Al} \to \mathrm{Al}$. Comme, en vertu des identités (3) et (5), pour toutes matrices d'octaves on a

$$\begin{aligned} \mathrm{Re}\,\mathrm{Tr}([A, X] \circ Y) &= \mathrm{Re}\,\mathrm{Tr}([A, X]\, Y) = \\ &= \mathrm{Re}\,\mathrm{Tr}(AX \cdot Y) - \mathrm{Re}\,\mathrm{Tr}(XA \cdot Y) = \\ &= \mathrm{Re}\,\mathrm{Tr}(Y \cdot AX) - \mathrm{Re}\,\mathrm{Tr}(X \cdot AY) = \end{aligned}$$

$$= \mathrm{Re}\,\mathrm{Tr}(YA \cdot X) - \mathrm{Re}\,\mathrm{Tr}(X \cdot AY) =$$
$$= -\mathrm{Re}\,\mathrm{Tr}(X \cdot [A,\ Y]) =$$
$$= -\mathrm{Re}\,\mathrm{Tr}(X \circ [A,\ Y]),$$

on obtient pour toute matrice antisymétrique A l'égalité

$$([A,\ X],\ Y) + (X,\ [A,\ Y]) = 0, \qquad X,\ Y \in \mathrm{Al}$$

qui exprime que l'opérateur ad A est antisymétrique pour le produit scalaire (2).

On sait par ailleurs que dans une algèbre alternative, l'associativité est réalisée pour des éléments a, b, c si certains d'entre eux sont égaux. Elle est aussi réalisée si l'un au moins des éléments a, b, c appartient à $\mathfrak{R}$, d'où il s'ensuit qu'elle est encore réalisée si certains des éléments a, b, c sont conjugués. En appliquant ce raisonnement aux éléments de la matrice $X(XX) - (XX)X$, où $X = (x_i^j)$ est une matrice d'octaves hermitienne d'ordre n, on trouve que dans les expressions $x_i^j(x_j^k x_k^l) - (x_i^j x_j^k)\, x_k^l$ des éléments non nuls de cette matrice, on peut étendre la sommation sur j et k aux seuls indices j et k, $j \neq k$, distincts de i et l. Donc, pour $n = 3$, on a nécessairement $i = l$, de sorte que quelle que soit la matrice $X \in \mathrm{Al}$, la matrice $X(XX) - (XX)X$ est obligatoirement diagonale, c'est-à-dire est de la forme

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}, \quad \text{où} \quad \alpha,\ \beta,\ \gamma \in \mathrm{Ca}.$$

Ceci étant, si la matrice X est de la forme (1), on aura alors les formules suivantes pour les octaves α, β, γ

$$\alpha = \xi_3(\xi_1\xi_2) + \bar{\xi}_2(\bar{\xi}_2\bar{\xi}_3) - (\xi_3\xi_1)\xi_2 - (\bar{\xi}_2\bar{\xi}_1)\bar{\xi}_3,$$
$$\beta = \xi_1(\xi_2\xi_3) + \bar{\xi}_3(\bar{\xi}_2\bar{\xi}_1) - (\xi_1\xi_2)\xi_3 - (\bar{\xi}_3\bar{\xi}_2)\bar{\xi}_1,$$
$$\gamma = \xi_2(\xi_3\xi_1) + \bar{\xi}_1(\bar{\xi}_3\bar{\xi}_2) - (\xi_2\xi_3)\xi_1 - (\bar{\xi}_1\bar{\xi}_3)\bar{\xi}_2,$$

d'où il résulte que $\alpha = \beta = \gamma$, puisque les associateurs sont antisymétriques.

Etant donné que pour toute matrice d'octaves A, on a $\mathrm{Tr}(A \cdot \alpha E) = \mathrm{Tr}\, A \cdot \alpha$, donc que $\mathrm{Tr}(A \cdot \alpha E) = 0$ si $\mathrm{Tr}\, A = 0$, on en déduit que pour toute matrice antihermitienne A d'ordre 3 de trace nulle et toute matrice $X \in \mathrm{Al}$ on a la relation

$$\mathrm{Tr}(A \cdot X(XX)) = \mathrm{Tr}(A \cdot (XX)X),$$

et, par suite (cf. formules (4) et (5)), la relation

$$\mathrm{Re}\,\mathrm{Tr}(AX \cdot XX) = \mathrm{Re}\,\mathrm{Tr}(XA \cdot XX),$$

qui exprime que

$$\operatorname{Re}\operatorname{Tr}((\operatorname{ad} A)X \cdot XX) = 0,$$

c'est-à-dire (puisque $(\operatorname{ad} A)X \in \mathrm{Al}$ et $XX = X \circ X$) que

$$((\operatorname{ad} A)X, X, X) = 0.$$

En polarisant cette identité en X, on obtient immédiatement que l'application linéaire $\operatorname{ad} A\colon \mathrm{Al} \to \mathrm{Al}$ satisfait l'identité (8).

Donc, l'application $\operatorname{ad} A$ est antisymétrique pour les deux produits scalaires (2) et (6). Par conséquent, c'est une dérivation de l'algèbre Al annulant la trace.

L'espace vectoriel des matrices d'octaves antihermitiennes d'ordre 3 de trace nulle sera désigné par $\mathcal{M}$. D'après ce qui a été prouvé, *pour toute matrice* $A \in \mathcal{M}$ *l'opérateur linéaire* $\operatorname{ad} A$ *est un élément de l'algèbre de Lie* $\mathfrak{f}_4$. L'application obtenue (qui est visiblement linéaire)

$$\operatorname{ad}\colon \mathcal{M} \to \mathfrak{f}_4$$

est *injective*. En effet, un calcul trivial montre que toute matrice de Al commute aux seules matrices scalaires de la forme aE, où $a \in \mathbb{R}$, or aE a une trace nulle uniquement pour $a = 0$. □

Les dérivations de l'algèbre Al de la forme $\operatorname{ad} A$, $A \in \mathcal{M}$, seront appelées *dérivations adjointes*.

Nous prêterons une attention particulière aux matrices de $\mathcal{M}$ dont les éléments diagonaux sont nuls. Ces matrices forment un sous-espace $\mathcal{M}^0$ de $\mathcal{M}$. Soient

$$Y_1(\eta) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \eta \\ 0 & -\bar{\eta} & 0 \end{pmatrix}, \quad Y_2(\eta) = \begin{pmatrix} 0 & 0 & -\bar{\eta} \\ 0 & 0 & 0 \\ \eta & 0 & 0 \end{pmatrix},$$

$$Y_3(\eta) = \begin{pmatrix} 0 & \eta & 0 \\ -\bar{\eta} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Toute matrice $A \in \mathcal{M}^0$ se représente de façon unique sous la forme

$$A = Y_1(\eta_1) + Y_2(\eta_2) + Y_3(\eta_3), \quad \eta_1, \eta_2, \eta_3 \in \mathrm{Ca},$$

d'où il vient, en particulier, que $\dim \mathcal{M}^0 = 24$.

Nous pouvons aborder maintenant la démonstration de l'important théorème de Freudenthal qui facilite sensiblement l'étude de l'algèbre Al et du groupe F_4. Comme toujours on désigne par $(F_4)_e$ la composante de l'unité du groupe F_4.

Proposition 1 (théorème de Freudenthal). *Pour tout élément* $X \in \mathrm{Al}$ *il existe un automorphisme* $\Phi \in (F_4)_e$ *tel que*

$$\Phi X = \lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3, \quad \text{où} \quad \lambda_1 \leqslant \lambda_2 \leqslant \lambda_3.$$

Les nombres $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$ *sont définis de façon unique par l'élément* X.

Deux éléments de l'algèbre Al *peuvent être associés par un automorphisme de* $(F_4)_e$ *si et seulement si les nombres* $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$ *qui leur correspondent sont confondus.*

Dans le langage de la théorie des groupes de représentations, cette proposition affirme que *chaque orbite du groupe* $(F_4)_e$ *dans l'algèbre* Al *contient une seule matrice diagonale* $\lambda_1 E_1 + \lambda_2 E_2 + \lambda_3 E_3$ *telle que* $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$. C'est sous cette forme que nous la prouverons.

Etant un sous-groupe fermé du groupe compact SO(26), le groupe $(F_4)_e$ est compact. Donc, chacune de ses orbites contient une matrice X de la forme (1) pour laquelle la somme $a_1^2 + a_2^2 + a_3^2$ prend son maximum (sur cette orbite). Il s'avère que *cette matrice est toujours diagonale.*

En effet, pour toute matrice $A \in \mathrm{Al}$ et tout $t \in \mathbb{R}$, la matrice $X_t = e^{t \operatorname{ad} A} X$ appartient à l'orbite de la matrice X (car $\operatorname{ad} A \in \mathfrak{f}_4$ et, par suite, $e^{t \operatorname{ad} A} \in (F_4)_e$), donc ses éléments diagonaux $a_1(t)$, $a_2(t)$, $a_3(t)$ vérifient l'inégalité $f(t) \leqslant f(0)$, où $f(t) = a_1(t)^2 + a_2(t)^2 + a_3(t)^2$. Mais nous savons déjà (cf. page 40) que la matrice X_t est solution de l'équation différentielle matricielle

$$\frac{dX_t}{dt} = (\operatorname{ad} A)\, X_t, \quad X_0 = X,$$

équation qui est équivalente à trois équations différentielles pour les octaves $\xi_1(t)$, $\xi_2(t)$ et $\xi_3(t)$ de la matrice X_t et à trois équations différentielles pour ses éléments numériques $a_1(t)$, $a_2(t)$, $a_3(t)$. Si $A = Y_1(\eta)$, un calcul simple utilisant la formule (2) de la leçon 14 montre que ces équations prennent la forme

$$\begin{aligned} &\frac{da_1(t)}{dt} = 0, && \frac{d\xi_1(t)}{dt} = (a_3(t) - a_1(t))\, \eta, \\ (9) \qquad &\frac{da_2(t)}{dt} = 2(\eta, \xi_1(t)), && \frac{d\xi_2(t)}{dt} = -\overline{\eta \xi_3(t)}, \\ &\frac{da_3(t)}{dt} = -2(\eta, \xi_1(t)), && \frac{d\xi_3(t)}{dt} = \overline{\xi_2(t)}\, \bar{\eta}. \end{aligned}$$

Donc, $a_1(t) = \text{const}$ et $a_2(t) + a_3(t) = \text{const}$. Par ailleurs, $f'(t) = 2(a_1'(t)\, a_1(t) + a_2'(t)\, a_2(t) + a_3'(t)\, a_3(t)) = 4(\eta, \xi_1(t))\, (a_2(t) - a_3(t))$, d'où, en vertu de l'égalité $f'(0) = 0$, on déduit que $a_2(0) = a_3(0)$ pour $(\eta, \xi_1(0)) \neq 0$. Donc, $a_1(t) = a_1(0)$ et $a_2(t) + a_3(t) = 2a_2(0)$ et, par suite, $f(t) = f(0) + 2(a_3(0) - a_3(t))^2$, ce qui en vertu de l'inégalité $f(t) \leqslant f(0)$ n'est possible

que pour $a_3(t) = a_3(0)$. Mais alors $a_3'(t) = 0$, et donc $(\eta, \xi_1(t)) = 0$, ce qui pour $t = 0$ contredit la condition $(\eta, \xi_1(0)) \neq 0$. Donc,

$$(\eta, \xi_1(0)) = 0$$

pour toute octave η, ce qui n'est possible que pour $\xi_1(0) = 0$.

On prouve de façon analogue que $\xi_2(0) = 0$ et $\xi_3(0) = 0$. Donc, la matrice $X = X_0$ est diagonale. □

Montrons maintenant que *toute matrice diagonale* $X = a_1E_1 + a_2E_2 + a_3E_3$ *peut être transformée par un automorphisme de* $(F_4)_c$ *en une matrice diagonale aussi telle que* $a_1 \leqslant a_2 \leqslant a_3$. Il est clair qu'il suffit de montrer que dans la matrice diagonale $X = a_1E_1 + a_2E_2 + a_3E_3$ on peut permuter deux éléments quelconques diagonaux, disons a_2 et a_3, par un automorphisme de $(F_4)_c$. Considérons à cet effet la matrice $X_t = e^{t \operatorname{ad} A}X$, où $A = Y_1(\eta)$. Des équations (9) il résulte aussitôt que pour cette matrice la fonction $a_2(t) - a_3(t)$ est solution de l'équation différentielle

$$\frac{d^2(a_2(t) - a_3(t))}{dt^2} = 4(\eta, \eta)(a_2(t) - a_3(t))$$

avec la condition initiale $a_2'(0) - a_3'(0) = 0$ (car $\xi_1(0) = 0$, puisque X est une matrice diagonale). Donc,

$$a_2(t) - a_3(t) = (a_2(0) - a_3(0)) \cos 2 | \eta | t.$$

En admettant pour simplifier que $| \eta | = 1$, on en déduit, en particulier, que

$$a_2\left(\frac{\pi}{2}\right) - a_3\left(\frac{\pi}{2}\right) = -(a_2(0) - a_3(0))$$

et, par suite (puisque $a_2(t) + a_3(t) =$ const), que $a_2\left(\frac{\pi}{2}\right) = a_3(0) = a_3$ et $a_3\left(\frac{\pi}{2}\right) = a_2(0) = a_2$. Donc, l'automorphisme $e^{\frac{\pi}{2} \operatorname{ad} Y_1(\eta)}$ (pour tout η tel que $| \eta | = 1$) permute bien a_2 et a_3. □

Ceci prouve entièrement la première assertion de la proposition 1. D'après cette assertion, chaque matrice $X \in \mathrm{Al}$ définit (pas forcément de façon unique) trois nombres $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$ tels que $\Phi X = \lambda_1E_1 + \lambda_2E_2 + \lambda_3E_3$ pour un automorphisme $\Phi \in (F_4)_c$.

Lemme 1. *On a les égalités suivantes*

$$\lambda_1 + \lambda_2 + \lambda_3 = (X, E),$$
$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = (X, X),$$
$$\lambda_1^3 + \lambda_2^3 + \lambda_3^3 = (X, X, X).$$

D é m o n s t r a t i o n. Pour la matrice diagonale ΦX, ces égalités sont évidentes (car $(X, E) = \operatorname{Tr} X$, $(X, X) = \operatorname{Tr}(X \circ X)$ et

$(X, X, X) = \mathrm{Tr}(X \circ X \circ X)$. Mais, comme $(\Phi X, E) = (\Phi X, \Phi E) = (X, E)$, $(\Phi X, \Phi X) = (X, X)$ et $(\Phi X, \Phi X, \Phi X) = (X, X, X)$, elles sont valables pour toute matrice $X \in \mathrm{Al}$. □

Nous pouvons maintenant achever la démonstration de la proposition 1.

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 1. Etant donné que nous avons déjà prouvé la première assertion de la proposition 1 et que la troisième assertion est une conséquence des deux premières, il nous suffit d'établir la seconde. Or, celle-ci résulte immédiatement du lemme 1, puisque d'après des formules classiques de théorie des polynômes symétriques les nombres λ_1, λ_2, λ_3 se définissent de façon unique à une permutation près par les nombres $\sigma_k = \lambda_1^k + \lambda_2^k + \lambda_3^k$, $k = 1, 2, 3$. □

Les nombres $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$ seront appelés *valeurs propres* de la matrice d'octaves X. A noter que leur somme est égale à la trace de la matrice X.

Le théorème de Freudenthal allège considérablement l'étude de l'algèbre Al. De ce théorème, il résulte par exemple que *les puissances de tout élément* $X \in \mathrm{Al}$ *sont associatives*, c'est-à-dire que les puissances n-ièmes de tout $X \in \mathrm{Al}$ sont les mêmes quelle que soit la place des parenthèses, puisque pour une matrice diagonale ceci est évident. (On démontre de façon analogue l'identité jordanienne (16) de la leçon 15 pour l'algèbre Al, puisque ceci est évident si l'un des facteurs est diagonal.)

Les puissances étant associatives, la puissance n-ième X^n d'un élément $X \in \mathrm{Al}$ est définie de façon unique pour tout $n \geqslant 0$. Donc, pour tout polynôme $p(T)$ à coefficients réels, est défini de façon unique un élément $p(X) \in \mathrm{Al}$. Si $p(X) = 0$, le polynôme $p(T)$ s'appelle *polynôme d'annulation* de l'élément X.

Le polynôme d'annulation de plus petit degré et de coefficient dominant égal à 1 s'appelle *polynôme minimal* de l'élément $X \in \mathrm{Al}$. Il est évident qu'il est unique et que pour tout automorphisme $\Phi\colon \mathrm{Al} \to \mathrm{Al}$ (qui n'est pas supposé préserver la trace) *les éléments* X *et* ΦX *admettent le même polynôme minimal.*

Si la matrice X est diagonale, ses éléments diagonaux sont racines de chaque polynôme d'annulation et, réciproquement, tout polynôme dont les racines sont ces éléments annule la matrice X. De là il s'ensuit que *le degré du polynôme minimal de toute matrice* $X \in \mathrm{Al}$ *est inférieur à* 3; *ce degré est égal à* 3 *si et seulement si les valeurs propres de la matrice* X *sont distinctes.*

Par ailleurs, on voit que dans le dernier cas le coefficient en T^2 du polynôme minimal est égal à $-(\lambda_1 + \lambda_2 + \lambda_3) = -\mathrm{Tr}\, X$. Le polynôme minimal étant invariant, on en déduit que tout automor-

phisme Φ: Al $\to$ Al vérifie l'égalité

$$\operatorname{Tr} \Phi X = \operatorname{Tr} X.$$

Pour des raisons de continuité, cette égalité est valable aussi dans le cas où les valeurs propres de l'élément X ne sont pas toutes distinctes. On a donc prouvé que *tout automorphisme de l'algèbre* Al *préserve la trace*, c'est-à-dire que le groupe F_4 est le groupe Aut Al des automorphismes de l'algèbre Al.

Du théorème de Freudenthal il résulte aussi immédiatement que *le groupe F_4 opère transitivement sur le plan projectif des octaves* Ca P² *des idempotents primitifs de l'algèbre* Al, c'est-à-dire que tout idempotent primitif peut être envoyé par un automorphisme dans un idempotent primitif. En effet, il est clair qu'une matrice diagonale est idempotente si et seulement si ses éléments diagonaux sont égaux à 0 ou à 1. Donc, les valeurs propres d'un idempotent primitif sont égales à 0, 0, 1 et, par suite, un tel idempotent est envoyé dans l'idempotent E_3 par un certain automorphisme. □

Cela signifie que *le plan projectif des octaves* Ca P² *est homéomorphe à la variété quotient F_4/K du groupe de Lie F_4 par son sous-groupe K* qui laisse invariant un idempotent primitif, pour fixer les idées, l'idempotent E_1.

Soit $\Phi \in K$. Comme $\Phi E_1 = E_1$, Φ envoie dans lui-même l'annulateur Ann E_1 de l'élément E_1, c'est-à-dire l'espace vectoriel des éléments $X \in$ Al tels que $X \circ E_1 = 0$. Mais il est aisé de voir que Ann E_1 est composé des éléments (1) pour lesquels $a_1 = 0$ et $\xi_2 = \xi_3 = 0$ et donc est somme directe du sous-espace à une dimension des matrices de la forme $a(E_2 + E_3)$ et du sous-espace $\mathrm{Ann}^0 E_1$ des matrices de la forme:

$$X = r(E_2 - E_3) + X_1(\rho), \quad \text{où} \quad r \in \mathbb{R}, \quad \rho \in \mathrm{Ca}.$$

L opérateur Φ étant orthogonal et $\Phi(E_2 + E_3) = \Phi(E - E_1) = E - E_1 = E_2 + E_3$, on en déduit que l'automorphisme Φ envoie dans lui-même le sous-espace $\mathrm{Ann}^0 E_1$, donc induit un opérateur orthogonal

$$\Phi^0 \colon \mathrm{Ann}^0 E_1 \to \mathrm{Ann}^0 E_1.$$

De façon analogue, l'automorphisme Φ envoie dans lui-même l'espace vectoriel $\mathcal{E}$ des éléments $X \in$ Al tels que $2E_1 \circ X = X$ et, par suite, induit un opérateur orthogonal

$$\Phi' \colon \mathcal{E} \to \mathcal{E}.$$

Ceci étant, d'après les formules (18), (19), (20) de la leçon précédente, l'espace vectoriel $\mathcal{E}$ est composé des matrices de la forme $X_2(\xi) + X_3(\eta)$, ξ, $\eta \in$ Ca, donc l'algèbre Al se décompose en somme directe de l'espace vectoriel $\mathcal{E}$, de l'espace $\mathrm{Ann}^0 E_1$ et des

deux espaces vectoriels à une dimension engendrés par les éléments E_1 et $E_2 + E_3$. Donc, *l'automorphisme* Φ *est défini de façon unique par les opérateurs* Φ^0 *et* Φ'.

On voit par ailleurs que l'espace $\mathscr{E}$ est canoniquement isomorphe à l'espace Ca^2 des couples d'octaves $\{\xi, \eta\}$ et l'espace $\mathrm{Ann}^0\,E_1$, à l'espace $\mathrm{Ca}^{\oplus}$ des couples $\{r, \rho\}$, où $r \in \mathbb{R}$, $\rho \in \mathrm{Ca}$. Donc, les opérateurs Φ' et Φ^0 peuvent être traités comme des opérateurs $\mathrm{Ca}^2 \to \mathrm{Ca}^2$ et $\mathrm{Ca}^{\oplus} \to \mathrm{Ca}^{\oplus}$ respectivement. Ceci étant, à l'élément $xu = \{-r\xi + \overline{\rho\eta}, r\eta + \overline{\xi\rho}\}$ de Ca^2, où $x = \{\xi, \eta\} \in \mathrm{Ca}^2$ et $u = \{r, \rho\} \in \mathrm{Ca}^{\oplus}$, sera associée la matrice

$$\begin{pmatrix} 0 & r\eta + \bar{\xi}\bar{\rho} & -r\bar{\xi} + \eta\rho \\ r\bar{\eta} + \rho\xi & 0 & 0 \\ -r\xi + \bar{\rho}\bar{\eta} & 0 & 0 \end{pmatrix} = $$

$$= \begin{pmatrix} 0 & \eta & \bar{\xi} \\ \bar{\eta} & 0 & 0 \\ \xi & 0 & 0 \end{pmatrix} \circ \begin{pmatrix} 0 & 0 & 0 \\ 0 & r & \rho \\ 0 & \bar{\rho} & -r \end{pmatrix},$$

de $\mathscr{E}$ qui est le produit jordanien des matrices de $\mathscr{E}$ et de $\mathrm{Ann}^0\,E_1$ associées aux éléments x et u. Les opérateurs Φ' et Φ^0 étant induits par l'automorphisme Φ de l'algèbre Al, on en déduit que ces opérateurs (traités comme des opérateurs $\mathrm{Ca}^2 \to \mathrm{Ca}^2$ et $\mathrm{Ca}^{\oplus} \to \mathrm{Ca}^{\oplus}$) satisfont l'identité

(10) $$\Phi'(xu) = \Phi'x \cdot \Phi^0 u, \quad x \in \mathrm{Ca}^2, \quad u \in \mathrm{Ca}^{\oplus}.$$

Donc, en vertu de la proposition 5 de la leçon 15, il existe un élément $a \in \mathrm{Spin}(9)$ et un seul tel que $\Phi' = a^R$ et $\Phi^0 = a^T$.

Ainsi, à tout automorphisme $\Phi \in K$ nous avons associé un élément $a \in \mathrm{Spin}(9)$. Il est clair que cette application est un homomorphisme. Bien plus, c'est un isomorphisme, car l'identité (10) est de toute évidence nécessaire et suffisante pour que l'application correspondante Φ soit un automorphisme de l'algèbre Al. On voit donc que *le sous-groupe* K *est canoniquement isomorphe au groupe spinoriel* Spin(9).

En identifiant le sous-groupe K au groupe Spin(9) au moyen de cet isomorphisme, on trouve en définitive que *le groupe* Spin(9) *est un sous-groupe du groupe* F_4, *et de plus la variété quotient* $F_4/\mathrm{Spin}(9)$ *est difféomorphe au plan projectif des octaves*:

(11) $$F_4/\mathrm{Spin}(9) \approx \mathrm{Ca}\,P^2.$$

De là il résulte, en particulier, que $\dim F_4 = \dim \mathrm{Spin}(9) + \dim \mathrm{Ca}\,P^2 = 36 + 16 = 52$.

Nous pouvons maintenant prouver pour le groupe F_4 une proposition totalement identique à la proposition 1 de la leçon 15 relative au groupe G_2:

Proposition 2. *Le groupe F_4 est connexe et simplement connexe. Tout groupe de Lie localement isomorphe au groupe F_4 est isomorphe à F_4.*

Démonstration. La première assertion résulte immédiatement de l'existence du difféomorphisme (11), puisque le groupe Spin(9) et le plan projectif Ca P² sont connexes et simplement connexes.

Pour prouver la deuxième assertion, il suffit de montrer que le centre du groupe F_4 est trivial (est composé seulement de l'automorphisme identique). Mais si un automorphisme $\Phi_0\colon \mathrm{Al} \to \mathrm{Al}$ appartient au centre du groupe F_4, alors vu qu'aucun idempotent primitif autre que E_1 ne laisse invariants les éléments du sous-groupe $K \approx \mathrm{Spin}(9)$, un raisonnement connu (cf. démonstration de la proposition 1 de la leçon 15) nous montre que $\Phi_0 \in K$, donc, que $\Phi_0 = \pm \mathrm{id}$ (car le centre du groupe Spin(9) est composé uniquement des éléments ± 1). Comme $(-E_1)^2 = E_1 \neq -E_1$, l'opérateur $-\mathrm{id}$ n'est pas un automorphisme. Donc, $\Phi_0 = \mathrm{id}$. □

Remarque 2. Si un automorphisme $\Phi\colon \mathrm{Al} \to \mathrm{Al}$ laisse invariant chaque idempotent E_i, $i = 1, 2, 3$, alors on voit aisément qu'à toute matrice $X_i(\xi)$, $i = 1, 2, 3$, il associe la matrice $X_i(\xi^{(i)})$, $i = 1, 2, 3$, où $\xi^{(i)}$ est une octave dépendant linéairement de l'octave ξ. On obtient ainsi trois opérateurs (visiblement orthogonaux) $\Phi\colon \xi \mapsto \xi^{(i)}$ de Ca dans Ca. Par un calcul direct utilisant la formule (20) de la leçon 15, on trouve que ces opérateurs satisfont l'identité

$$\Phi_1\xi\cdot\Phi_2\eta = \Phi_3(\overline{\xi\eta}), \qquad \xi, \eta \in \mathrm{Ca}.$$

Donc, en vertu de la proposition 3 de la leçon 15, il existe un élément $a \in \mathrm{Spin}(8)$ et un seul tel que $\Phi_1 = a^L$, $\Phi_2 = a^R$ et $\Phi_3 \circ \sigma = a^T$, où $\sigma\colon \xi \mapsto \bar\xi$ est la conjugaison sur l'algèbre Ca. La correspondance $\Phi \mapsto a$ étant de toute évidence un isomorphisme, ceci prouve que *le groupe* Aut^0 Al *des automorphismes de l'algèbre* Al *laissant invariants les idempotents* E_i, $i = 1, 2, 3$, *est isomorphe au groupe* Spin(8).

Le groupe de Lie $F_4 = \mathrm{Aut}$ Al étant simplement connexe, sa structure algébrique est entièrement définie par celle de l'algèbre de Lie $\mathfrak{f}_4 = \mathrm{Der}$ Al. Il nous reste donc à étudier cette algèbre de Lie.

Considérons à cet effet la sous-algèbre Der^0 Al de $\mathfrak{f}_4$ composée des dérivations annulant chaque idempotent E_i, $i = 1, 2, 3$. Il est clair que cette sous-algèbre est l'algèbre de Lie du sous-groupe Aut^0 Al (cf. remarque 2) et, par suite, est isomorphe à l'algèbre de Lie $\mathfrak{l}(\mathrm{Spin}\,(8)) \approx \mathfrak{so}(8)$. Du reste, on peut directement et sans peine établir cet isomorphisme. En effet, si $DE_i = 0$ pour $i =$

$= 1, 2, 3$, alors ainsi qu'il résulte de la formule (19) de la leçon 15, l'élément $DX_i(\xi)$, $\forall i = 1, 2, 3$, est de la forme $X_i(A_i\xi)$, où A_i est un opérateur linéaire $\mathbb{C}a \to \mathbb{C}a$. Les opérateurs A_i sont antisymétriques et satisfont (ceci résulte immédiatement pour $i = 1$ des formules (20) de la leçon 15) l'identité

$$A_1\xi\cdot\eta + \xi\cdot A_2\eta = A_3\,(\overline{\xi\eta}). \tag{12}$$

Donc (cf. remarque 2 de la leçon 15), $A_2 = (A_1^{\lambda^{-1}})^\rho$ et $A_3 = A_1^{\lambda^{-1}} \circ \sigma$, où σ: $\mathbb{C}a \to \mathbb{C}a$ est la conjugaison. Comme, réciproquement, tous opérateurs A_1, A_2, A_3 satisfaisant l'identité (12) sont de toute évidence associés à une dérivation $D \in \mathrm{Der}^0\,\mathrm{Al}$, ceci prouve que la correspondance $D \mapsto A_1$ est un isomorphisme. □

Nous désignerons la dérivation de $\mathrm{Der}^0\,\mathrm{Al}$ associée à un opérateur $A \in \mathfrak{so}(8)$ par $\varkappa A$. Donc, $D = \varkappa A_1$ si et seulement si $DE_i = 0$ pour tout $i = 1, 2, 3$ et $DX_1(\xi) = X_1(A\xi)$ pour tout $\xi \in \mathbb{C}a$.

Soit maintenant D une dérivation quelconque de l'algèbre Al, et soit

$$DE_i = a_{1i}E_1 + a_{2i}E_2 + a_{3i}E_3 + X_1(\xi_{1i}) + X_2(\xi_{2i}) + X_3(\xi_{3i}).$$

Comme $E_i^2 = E_i$, il vient $2DE_i \circ E_i = DE_i$, d'où, en vertu des formules (18), (19), (20) de la leçon 15, il résulte immédiatement que $a_{ij} = 0$ pour tous i, j et $\xi_{ii} = 0$. Par ailleurs, comme $E_i \circ E_j = 0$ pour $i \neq j$, on a $DE_i \circ E_j + E_i \circ DE_j = 0$, d'où il s'ensuit que $\xi_{i,i+1} + \xi_{i,i+2} = 0$ pour tout $i = 1, 2, 3$. Donc, en posant $\xi_i = \xi_{i,i+1}$, on trouve que

$$DE_i = X_{i+1}(\xi_{i+1}) - X_{i+2}(\xi_{i+2}), \quad i = 1, 2, 3 \bmod 3. \tag{13}$$

D'autre part, un calcul immédiat montre que la dérivation adjointe $D = \mathrm{ad}\,A$ associée à la matrice $A = Y_1(\xi_1) + Y_2(\xi_2) + Y_3(\xi_3)$ de $\mathcal{M}^0$ est justiciable des mêmes formules (13). Ceci signifie que pour toute dérivation $D \in \mathrm{Der}\,\mathrm{Al}$, il existe une dérivation adjointe $\mathrm{ad}\,A$, $A \in \mathcal{M}^0$, et une seule telle que $D - \mathrm{ad}\,A \in \mathrm{Der}^0\,\mathrm{Al} = \varkappa(\mathfrak{so}(8))$, c'est-à-dire que

$$\mathfrak{f}_4 = \varkappa\,(\mathfrak{so}(8)) \oplus \mathrm{ad}\,\mathcal{M}^0. \tag{14}$$

Les applications $\varkappa$: $\mathfrak{so}(8) \to \mathfrak{f}_4$ et ad: $\mathcal{M}^0 \to \mathfrak{f}_4$ étant des monomorphismes, ceci prouve que

$$\mathfrak{f}_4 \approx \mathfrak{so}\,(8) \oplus \mathcal{M}^0.$$

Comme $\dim \mathfrak{so}(8) = 28$ et $\dim \mathcal{M}^0 = 24$, ceci prouve encore que $\dim \mathfrak{f}_4 = 52$.

En vertu de la décomposition (14), pour déterminer entièrement la structure algébrique de l'algèbre de Lie $\mathfrak{f}_4$, il suffit de calculer tous les crochets de Lie des éléments $\varkappa A$ et $\mathrm{ad}\,Y_i(\xi)$, où $A \in \mathfrak{so}(8)$,

$\xi \in \mathrm{Ca}$, $i = 1, 2, 3$. Il s'avère que

$$[\varkappa A, \varkappa B] = \varkappa [A, B], \tag{15}$$

$$[\varkappa A, \operatorname{ad} Y_i(\xi)] = \operatorname{ad} Y_i(A\xi), \tag{16}$$

$$[\operatorname{ad} Y_i(\xi), \operatorname{ad} Y_j(\eta)] = \begin{cases} \varkappa C^{(i)}_{\xi,\eta} & \text{pour } j = i, \\ \operatorname{ad} Y_{i+2}(-\overline{\xi\eta}) & \text{pour } j = i+1, \end{cases} \tag{17}$$

où A et B sont des éléments quelconques de l'algèbre de Lie $\mathfrak{so}(8)$ (traités comme des opérateurs antisymétriques $\mathrm{Ca} \to \mathrm{Ca}$), ξ, η des octaves arbitraires, $C^{(i)}_{\xi,\eta}$ des opérateurs $\mathrm{Ca} \to \mathrm{Ca}$, définis respectivement par les formules

$$\begin{aligned} C^{(1)}_{\xi,\eta}&: \ \zeta \to 4(\xi, \eta)\eta - 4(\eta, \xi)\xi, \\ C^{(2)}_{\xi,\eta}&: \ \zeta \mapsto \zeta\xi\cdot\bar{\eta} - \zeta\eta\cdot\bar{\xi}, \qquad\qquad \zeta \in \mathrm{Ca}. \\ C^{(3)}_{\xi,\eta}&: \ \zeta \mapsto \bar{\eta}\cdot\xi\zeta - \bar{\xi}\cdot\eta\zeta, \end{aligned}$$

En effet, la formule (15) revient à affirmer que l'application $\varkappa$ est un homomorphisme, et les formules (16) et (17) se vérifient par un calcul direct. Par exemple, il est aisé de voir que

$$[Y_1(\xi), E_1] = 0, \quad [Y_1(\xi), E_2] = -X_1(\xi), \quad [Y_1(\xi), E_3] = X_1(\xi),$$

c'est-à-dire que $(\operatorname{ad} Y_1(\xi))E_i = \varepsilon_i X_1(\xi)$, où $\varepsilon_i = 0, -1, 1$ pour $i = 1, 2, 3$. Donc (puisque par définition $(\varkappa A)E_i = 0$ pour tout $i = 1, 2, 3$),

$$\begin{aligned} [\varkappa A, \operatorname{ad} Y_1(\xi)] E_i &= (\varkappa A)(\operatorname{ad} Y_1(\xi))E_i = \\ &= \varepsilon_i X_1(A\xi) = (\operatorname{ad} Y_1(A\xi)) E_i, \end{aligned}$$

et donc la dérivation $D = [\varkappa A, \operatorname{ad} Y_1(\xi)] - \operatorname{ad} Y_1(A\xi)$ appartient à $\mathrm{Der}^0 \mathrm{Al}$. Mais, un calcul immédiat montre que

$$[Y_1(\xi), X_1(\zeta)] = 2(\xi, \zeta)(E_2 - E_3),$$

donc pour D on a la formule suivante

$$\begin{aligned} DX_1(\zeta) &= (\varkappa A)(2(\xi, \zeta)(E_2 - E_3)) - \\ &\quad - [Y_1(\xi), X_1(A\zeta)] - [Y_1(A\xi), X_1(\zeta)] = \\ &= -2[(\xi, A\xi) + (A\xi, \zeta)](E_2 - E_3) = 0, \end{aligned}$$

car l'opérateur A est antisymétrique par hypothèse. Donc, $D = 0$, ce qui prouve la formule (16) pour $i = 1$. De façon analogue

$$\begin{aligned} [\operatorname{ad} Y_1(\xi), \operatorname{ad} Y_1(\eta)]E_i &= \\ &= \varepsilon_i(\operatorname{ad} Y_1(\xi)) X_1(\eta) - \varepsilon_i(\operatorname{ad} Y_1(\eta)) X_1(\xi) = \\ &= 2\varepsilon_i((\xi, \eta) - (\eta, \xi))(E_2 - E_3) = 0 \end{aligned}$$

et

$[\text{ad}\, Y_1(\xi), \text{ad}\, Y_1(\eta)]X_1(\zeta) =$

$$= 2\,((\eta, \zeta)\,\text{ad}\, Y_1(\xi) - (\xi, \zeta)\,\text{ad}\, Y_1(\eta))\,(E_2 - E_3) =$$

$$= 4\,(-(\eta, \zeta)\, X_1(\xi) + (\xi, \zeta) X_1(\eta)),$$

ce qui prouve la formule (17) pour $i = j = 1$. Les autres formules (16) et (17) se vérifient de façon analogue.

Les résultats acquis nous permettent de prouver pour l'algèbre de Lie $\mathfrak{f}_4$ une proposition analogue à la proposition 1 de la leçon 14.

Soit $\mathfrak{h}$ le sous-espace à quatre dimensions de l'algèbre de Lie $\mathfrak{f}_4$, composé des combinaisons linéaires des éléments $\varkappa E_{[1,8]}$, $\varkappa E_{[2,7]}$, $\varkappa E_{[3,6]}$ et $\varkappa E_{[4,5]}$. Comme ces éléments commutent l'un à l'autre, il vient que $[H_1, H_2] = 0$ pour tous éléments $H_1, H_2 \in \mathfrak{h}$. Soit e_1, e_2, e_3, e_4 la base de l'espace dual $\mathfrak{h}^*$, duale de la base $\varkappa E_{[1,8]}$, $\varkappa E_{[2,7]}$, $\varkappa E_{[3,6]}$, $-\varkappa E_{[4,5]}$ de l'espace $\mathfrak{h}$. Munissons $\mathfrak{h}^*$ d'une structure euclidienne pour laquelle la base e_1, e_2, e_3, e_4 est orthonormée et appelons *configuration* F_4 l'ensemble des vecteurs de la forme

$$\pm e_p, \quad \pm e_p \pm e_q \quad (p \neq q),$$

$$\frac{1}{2}(\pm e_1 \pm e_2 \pm e_3 \pm e_4),$$

(en combinant ces signes on obtient 48 vecteurs).

En vertu de l'identification $\mathfrak{h}^* = \mathfrak{h}$ induite par la structure euclidienne, on peut traiter les vecteurs $\alpha \in F_4$ comme des éléments de $\mathfrak{h}$. Suivant l'exemple de l'algèbre de Lie $\mathfrak{g}_2^{\mathbb{C}}$ (cf. leçon 14), posons

$$H_\alpha = \frac{2\alpha}{|\alpha|^2} \text{ pour tout } \alpha \in F_4.$$

Pour éviter toute confusion avec les éléments de Ca, désignons l'unité imaginaire du corps $\mathbb{C}$ par $\sqrt{-1}$ et définissons les éléments X_α, $\alpha = \pm e_p \pm e_q$ de l'algèbre de Lie complexifiée $\mathfrak{f}_4^{\mathbb{C}} = \mathfrak{f}_4 \otimes \mathbb{C}$ en posant

$$X_{e_1 \pm e_2} = \varkappa\,(E_{[7,8]} \pm E_{[1,2]}) + \sqrt{-1} \cdot \varkappa\,(E_{[7,1]} \pm E_{[2,8]}),$$

$$X_{-e_1 \pm e_2} = \varkappa\,(-E_{[7,8]} \pm E_{[1,2]}) + \sqrt{-1} \cdot \varkappa\,(E_{[7,1]} \mp E_{[2,8]}),$$

et en admettant que les autres vecteurs $X_{\pm e_p \pm e_q}$ sont définis par les formules obtenues en changeant les indices (7, 8) par les indices (2, 7), (3, 6), (5, 4) lorsque $p = 2, 3, 4$, et respectivement, les indices (1, 2) par les indices (1, 8), (3, 6), (5, 4) lorsque $q = 1, 3, 4$.

Pour $\alpha = \pm e_p$, nous posons

$$X_{\pm e_p} = \text{ad}\, Y_1(\xi_p) \pm \sqrt{-1}\ \text{ad}\, Y_1(\eta_p),$$

où

$$\xi_1 = 1, \qquad \xi_2 = i, \qquad \xi_3 = j, \qquad \xi_4 = k,$$
$$\eta_1 = -h, \quad \eta_2 = -g, \quad \eta_3 = -f, \quad \eta_4 = -e,$$

et pour $\alpha = \frac{1}{2}(\pm e_1 \pm e_2 \pm e_3 \pm e_4)$,

$$X_\alpha = \begin{cases} \operatorname{ad} Y_2(\xi_p) + \sqrt{-1}\ \operatorname{ad} Y_2(\eta_p), \\ \operatorname{ad} Y_3(\xi_p) + \sqrt{-1}\ \operatorname{ad} Y_3(\eta_p), \end{cases}$$

si

$$\alpha = \begin{cases} \pm \frac{1}{2}(-e_1 + e_2 + e_3 + e_4), \\ \pm \frac{1}{2}(-e_1 + e_2 - e_3 - e_4), \\ \pm \frac{1}{2}(-e_1 - e_2 + e_3 - e_4), \\ \pm \frac{1}{2}(-e_1 - e_2 - e_3 + e_4), \end{cases}$$

ou respectivement

$$\alpha = \begin{cases} \pm \frac{1}{2}(-e_1 - e_2 - e_3 - e_4), \\ \pm \frac{1}{2}(e_1 + e_2 - e_3 - e_4), \\ \pm \frac{1}{2}(e_1 - e_2 + e_3 - e_4), \\ \pm \frac{1}{2}(e_1 - e_2 - e_3 + e_4), \end{cases}$$

où p est le numéro du vecteur α correspondant dans ces formules.

Il s'avère alors que *la proposition* 1 *de la leçon* 14 (*de même que la formule* (17) *de la même leçon*) *est valable pour l'algèbre* $\mathfrak{f}_4^{\mathbb{C}}$ (relativement bien sûr à la configuration F_4).

On peut prouver cette proposition en choisissant une base convenable dans $\mathfrak{f}_4^{\mathbb{C}}$ et en reproduisant pas à pas les calculs qui nous ont permis d'établir la proposition 1 et la formule (17) pour l'algèbre de Lie $\mathfrak{g}_2^{\mathbb{C}}$. A noter que les calculs seront bien plus longs ici. Au prochain semestre nous développerons une théorie générale qui permettra de réduire ce travail, c'est pourquoi nous ajournons cette démonstration.

D'après ce qui précède, il n'est pas surprenant qu'une proposition analogue à la proposition 2 de la leçon 14 soit valable pour l'algèbre $\mathfrak{f}_4^{\mathbb{C}}$. Pour formuler cette proposition, il suffit de considérer

que dans les formules (23) de la leçon 14 les indices p et q varient entre 1 et 4, et de traiter les n_{pq} comme les éléments de la matrice

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

La démonstration de cette proposition repose sur les mêmes considérations, mais implique plus de travail. Le rôle des vecteurs f_1, e_2 sera tenu par les vecteurs

$$\frac{1}{2}(e_1-e_2-e_3-e_4),\ e_4,\ e_3-e_4,\ e_2-e_3.$$

Le détail des calculs est laissé au soin du lecteur.

LEÇON 17

Algèbres de Lie résolubles.— Radical d'une algèbre de Lie. —Algèbres de Lie abéliennes.— Centre d'une algèbre de Lie. —Algèbres de Lie nilpotentes.— Nilradical d'une algèbre de Lie.— Algèbres de Lie nilpotentes linéaires.— Théorème d'Engel.— Critères de nilpotence.— Algèbres de Lie irréductibles linéaires.— Algèbres de Lie réductives.— Algèbres de Lie résolubles linéaires.— Radical nilpotent d'une algèbre de Lie.

Considérons maintenant la théorie générale des algèbres de Lie dans le but de prouver le théorème d'Ado et de combler ainsi la lacune laissée dans la démonstration du théorème de Cartan dans la leçon 10.

La démonsration du théorème d'Ado repose sur la théorie structurelle des algèbres de Lie qui est assez élaborée et qui présente un grand intérêt en soi. Malheureusement nous ne pourrons qu'effleurer cette théorie.

Sauf mention expresse du contraire, dans toute la suite le corps de base $\mathcal{K}$ est un corps arbitraire de caractéristique zéro. Toutes les algèbres de Lie sur $\mathcal{K}$ sont supposées de dimension finie.

Soit $\mathfrak{g}$ une algèbre de Lie sur $\mathcal{K}$. Etant donnés deux sous-espaces $\mathfrak{a}$ et $\mathfrak{b}$ de $\mathfrak{g}$, on désignera par $[\mathfrak{a}, \mathfrak{b}]$ le sous-espace engendré par les éléments $[a, b]$, où $a \in \mathfrak{a}$, $b \in \mathfrak{b}$. Dans ces notations la propriété du sous-espace $\mathfrak{a}$ d'être une sous-algèbre équivaut à l'inclusion $[\mathfrak{a}, \mathfrak{a}] \subset \mathfrak{a}$, et la propriété d'être un idéal, à l'inclusion $[\mathfrak{a}, \mathfrak{g}] \subset \subset \mathfrak{a}$. Comme $[[\mathfrak{a}, \mathfrak{b}], \mathfrak{c}] \subset [[\mathfrak{b}, \mathfrak{c}], \mathfrak{a}] + [[\mathfrak{c}, \mathfrak{a}], \mathfrak{b}]$ d'après l'identité de Jacobi, on en déduit, en particulier, que *si* $\mathfrak{a}$ *et* $\mathfrak{b}$ *sont des idéaux, il en est de même de* $[\mathfrak{a}, \mathfrak{b}]$.

Donc, les formules

$$\mathfrak{g}^{(1)} = \mathfrak{g}, \quad \mathfrak{g}^{(2)} = [\mathfrak{g}, \mathfrak{g}], \ \ldots, \ \mathfrak{g}^{(k)} = [\mathfrak{g}^{(k-1)}, \mathfrak{g}^{(k-1)}], \ldots$$

définissent dans l'algèbre de Lie $\mathfrak{g}$ une suite décroissante d'idéaux

$$\mathfrak{g}=\mathfrak{g}^{(1)}\supset\mathfrak{g}^{(2)}\supset\ldots\supset\mathfrak{g}^{(k)}\supset\ \ldots$$

A noter que $\mathfrak{g}^{(k)(l)}=\mathfrak{g}^{(k+l-1)}$.

Certains auteurs désignent l'idéal $\mathfrak{g}^{(k)}$ par $\mathfrak{g}^{(k-1)}$.

Définition 1. On dit qu'une algèbre de Lie $\mathfrak{g}$ est *résoluble* s'il existe un $k\geqslant 0$ tel que $\mathfrak{g}^{(k)}=0$.

Soit $n=\dim\mathfrak{g}$. La suite décroissante de sous-espaces

$$\mathfrak{g}=\mathfrak{g}_0\supset\mathfrak{g}_1\supset\ldots\supset\mathfrak{g}_i\supset\ldots\supset\mathfrak{g}_n=0$$

s'appelle *drapeau* si la dimension de chaque sous-espace est inférieure d'une unité à celle du précédent, c'est-à-dire si $\dim\mathfrak{g}_i=$ $=n-i$ pour tout $i=0,1,\ldots,n$. Un drapeau composé de sous-algèbres s'appelle *drapeau de sous-algèbres*.

Proposition 1. *Une algèbre de Lie $\mathfrak{g}$ est résoluble si et seulement si elle contient un drapeau de sous-algèbres*

$$\mathfrak{g}=\mathfrak{g}_0\supset\mathfrak{g}_1\supset\ldots\supset\mathfrak{g}_i\supset\ldots\supset\mathfrak{g}_n=0$$

dans lequel chaque sous-algèbre $\mathfrak{g}_i$, $i=1,\ldots,n$, est un idéal de la sous-algèbre précédente $\mathfrak{g}_{i-1}$ (satisfait la relation $[\mathfrak{g}_{i-1},\mathfrak{g}_i]\subset\mathfrak{g}_i$).

Démonstration. Par hypothèse $\dim\mathfrak{g}_{i-1}=\dim\mathfrak{g}_i+1$. Donc, tout élément de $\mathfrak{g}_{i-1}$ est de la forme $x+\lambda e$, où $x\in\mathfrak{g}_i$, $\lambda\in K$, et e un élément fixe. Comme

$$[x+\lambda e,\ y+\mu e]=[x,y]+\lambda[x,e]-\mu[y,e],$$

et $[x,y]$, $[x,e]$, $[y,e]\in\mathfrak{g}_i$ (car $\mathfrak{g}_i$ est un idéal de $\mathfrak{g}_{i-1}$), il en résulte que $[\mathfrak{g}_{i-1},\mathfrak{g}_{i-1}]\subset\mathfrak{g}_i$. Donc, si $\mathfrak{g}^{(i)}\subset\mathfrak{g}_{i-1}$, alors $\mathfrak{g}^{(i+1)}=[\mathfrak{g}^{(i)},\mathfrak{g}^{(i)}]\subset[\mathfrak{g}_{i-1},\mathfrak{g}_{i-1}]\subset\mathfrak{g}_i$. Etant vraie pour $i=1$, l'inclusion $\mathfrak{g}^{(i)}\subset$ $\subset\mathfrak{g}_{i-1}$ l'est pour tout $i=1,\ldots,n+1$. En particulier, $\mathfrak{g}^{(n+1)}\subset$ $\subset\mathfrak{g}_n=0$, c'est-à-dire que $\mathfrak{g}^{(n+1)}=0$.

Réciproquement, s'il existe un $k\geqslant 0$ tel que $\mathfrak{g}^{(k)}=0$ (et $\mathfrak{g}^{(k-1)}\neq 0$), alors dans la suite d'idéaux

$$\mathfrak{g}=\mathfrak{g}^{(0)}\supset\mathfrak{g}^{(1)}\supset\ldots\supset\mathfrak{g}^{(k)}=0$$

les inclusions sont toutes strictes, et, par conséquent, cette suite peut être incluse dans un drapeau

$$\mathfrak{g}=\mathfrak{g}_0\supset\mathfrak{g}_1\supset\ldots\supset\mathfrak{g}_n=0.$$

Soit $i=0,1,\ldots,n$, et soit a le plus grand indice tel que $\mathfrak{g}_i\subset\mathfrak{g}^{(a)}$. Si $\mathfrak{g}_i\neq\mathfrak{g}^{(a)}$, alors $\mathfrak{g}_{i-1}\subset\mathfrak{g}^{(a)}$ et donc $[\mathfrak{g}_{i-1},\mathfrak{g}_i]\subset[\mathfrak{g}^{(a)},\mathfrak{g}^{(a)}]=$ $=\mathfrak{g}^{(a+1)}\subset\mathfrak{g}_i$. Si $\mathfrak{g}_i=\mathfrak{g}^{(a)}$, alors $\mathfrak{g}_{-i1}\subset\mathfrak{g}^{(a-1)}$ et, par suite, $[\mathfrak{g}_{i-1},\mathfrak{g}_i]\subset[\mathfrak{g}^{(a-1)},\mathfrak{g}^{(a)}]\subset\mathfrak{g}^{(a)}=\mathfrak{g}_i$. Donc, dans tous les cas, on a $[\mathfrak{g}_{i-1},\mathfrak{g}_i]\subset\mathfrak{g}_i$, d'où il s'ensuit que le drapeau considéré est

un drapeau de sous-algèbres (car $[\mathfrak{g}_i, \mathfrak{g}_i] \subset [\mathfrak{g}_{i-1}, \mathfrak{g}_i] \subset \mathfrak{g}_i$) dans lequel chaque sous-algèbre est un idéal de l'algèbre précédente. □

Les idéaux $\mathfrak{g}^{(k)}$ se transforment visiblement en les idéaux $\mathfrak{h}^{(k)}$ par tout épimorphisme $\mathfrak{g} \to \mathfrak{h}$. De façon analogue, si $\mathfrak{h} \subset \mathfrak{g}$, alors $\mathfrak{h}^{(k)} \subset \mathfrak{g}^{(k)} \cap \mathfrak{h}$. Donc, *toute algèbre quotient et toute sous-algèbre d'une algèbre de Lie résoluble sont résolubles.* Par ailleurs, il est aisé de voir qu'*une algèbre $\mathfrak{g}$ est résoluble si elle contient un idéal résoluble $\mathfrak{h}$ tel que l'algèbre quotient $\mathfrak{g}/\mathfrak{h}$ soit résoluble.* En effet, si $(\mathfrak{g}/\mathfrak{h})^{(k)} = 0$, alors $\mathfrak{g}^{(k)} = \mathfrak{h}$ et, par suite, si $\mathfrak{h}^{(l)} = 0$, alors $\mathfrak{g}^{(k+l-1)} = \mathfrak{g}^{(k)(l)} = 0$. □

La somme $\mathfrak{a} + \mathfrak{b}$ de deux idéaux $\mathfrak{a}$ et $\mathfrak{b}$ est visiblement un idéal et, de plus, d'après le premier théorème des isomorphismes, l'algèbre quotient $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b}$ est isomorphe à l'algèbre quotient $\mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$ et donc est résoluble si l'idéal $\mathfrak{a}$ l'est. Par conséquent, si l'idéal $\mathfrak{b}$ est résoluble aussi, il en sera de même de l'idéal $\mathfrak{a} + \mathfrak{b}$. Donc, *la somme $\mathfrak{a} + \mathfrak{b}$ de deux idéaux résolubles est un idéal résoluble.* C'est pourquoi *dans toute algèbre de Lie $\mathfrak{g}$ de dimension finie, il existe un plus grand idéal résoluble* $\mathfrak{r}$ qui contient tous les autres idéaux résolubles de $\mathfrak{g}$: cet idéal est la somme de tous les idéaux résolubles de $\mathfrak{g}$.

Définition 2. L'idéal $\mathfrak{r}$ s'appelle *le radical* de l'algèbre de Lie $\mathfrak{g}$. Si $\mathfrak{r} = 0$, l'algèbre $\mathfrak{g}$ est dite *semi-simple.*

A noter que tout homomorphisme $\mathfrak{g} \to \mathfrak{h}$ envoie le radical de l'algèbre $\mathfrak{g}$ dans celui de l'algèbre $\mathfrak{h}$.

L'épimorphisme canonique $\mathfrak{g} \to \mathfrak{g}/\mathfrak{r}$ établit une correspondance biunivoque entre les idéaux $\mathfrak{a}$ de l'algèbre $\mathfrak{g}$ contenant l'idéal $\mathfrak{r}$, et les idéaux $\mathfrak{b}$ de l'algèbre $\mathfrak{g}/\mathfrak{r}$, et de plus l'idéal $\mathfrak{b}$ associé à l'idéal $\mathfrak{a}$ est isomorphe à l'algèbre quotient $\mathfrak{a}/\mathfrak{r}$ et, par suite, est résoluble si et seulement si l'idéal $\mathfrak{a}$ l'est. Or, $\mathfrak{r}$ étant le radical maximal de $\mathfrak{g}$, l'idéal $\mathfrak{a} \supset \mathfrak{r}$ est résoluble si et seulement si $\mathfrak{a} = \mathfrak{r}$. Ceci prouve que *l'algèbre quotient* $\mathfrak{a}/\mathfrak{r}$ ne contient pas d'idéaux résolubles non nuls, c'est-à-dire qu'elle *est semi-simple.*

Définition 3. On dit qu'une algèbre de Lie $\mathfrak{g}$ est *abélienne* si $\mathfrak{g}^{(2)} = 0$, c'est-à-dire si $[x, y] = 0$ quels que soient $x, y \in \mathfrak{g}$.

Si une algèbre de Lie $\mathfrak{g}$ n'est pas semi-simple, c'est-à-dire si son radical $\mathfrak{r}$ est non nul, et si k est le plus petit indice tel que $\mathfrak{r}^{(k)} = 0$, alors l'idéal $\mathfrak{a} = \mathfrak{r}^{(k-1)}$ (qui est visiblement un idéal dans $\mathfrak{g}$ aussi) est non nul et abélien ($[\mathfrak{a}, \mathfrak{a}] = [\mathfrak{r}^{(k-1)}, \mathfrak{r}^{(k-1)}] = \mathfrak{r}^{(k)} = 0$). Réciproquement, si une algèbre de Lie $\mathfrak{g}$ contient un idéal abélien $\mathfrak{a} \neq 0$ (donc résoluble), alors $\mathfrak{r} \neq 0$ et, par suite, l'algèbre de Lie $\mathfrak{g}$ n'est pas semi-simple. Donc, *une algèbre de Lie $\mathfrak{g}$ est semi-simple si et seulement si elle ne contient pas d'idéaux abéliens non nuls.*

On appelle *centre* d'une algèbre de Lie $\mathfrak{g}$ son annulateur au sens de la théorie générale des algèbres, c'est-à-dire le plus grand sous-espace $\mathfrak{z} \subset \mathfrak{g}$ tel que $[\mathfrak{z}, \mathfrak{g}] = 0$. On vérifie immédiatement que *le centre est un idéal.*

Une algèbre $\mathfrak{g}$ est abélienne si et seulement si $\mathfrak{z} = \mathfrak{g}$.

Le centre étant un idéal abélien, *le centre d'une algèbre semi-simple est nul.*

Outre les idéaux $\mathfrak{g}^{(k)}$, on peut considérer aussi les idéaux $\mathfrak{g}^k$ définis par la formule récurrentielle

$$\mathfrak{g}^k = [\mathfrak{g}, \mathfrak{g}^{k-1}], \quad (\mathfrak{g}^1 = \mathfrak{g}).$$

On remarquera que $\mathfrak{g}^2 = \mathfrak{g}^{(2)}$.

Définition 4. Une algèbre de Lie $\mathfrak{g}$ est *nilpotente* s'il existe un $k \geqslant 1$ tel que $\mathfrak{g}^k = 0$.

Comme $\mathfrak{g}^2 = \mathfrak{g}^{(2)}$, on voit, en particulier, que *toute algèbre de Lie abélienne est nilpotente.*

Une récurrence sur i nous permet d'établir immédiatement que $\mathfrak{g}^{(i)} \subset \mathfrak{g}^i$ pour tout i. Donc, *toute algèbre de Lie nilpotente est résoluble.*

Si $\mathfrak{g}^k = 0$ et $\mathfrak{g}^{k-1} \neq 0$, alors l'idéal non nul $\mathfrak{a} = \mathfrak{g}^{k-1}$ est tel que $[\mathfrak{a}, \mathfrak{g}] = 0$ et, par suite, appartient au centre $\mathfrak{z}$ de l'algèbre $\mathfrak{g}$. Donc, *le centre $\mathfrak{z}$ d'une algèbre de Lie nilpotente est non nul.*

A noter que le centre d'une algèbre de Lie résoluble peut très bien être nul.

Proposition 2. *Une algèbre de Lie $\mathfrak{g}$ est nilpotente si et seulement s'il existe un drapeau de sous-algèbres*

$$\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \ldots \supset \mathfrak{g}_i \supset \ldots \supset \mathfrak{g}_n = 0,$$

tel que pour tout $i = 1, \ldots, n$ l'on ait l'inclusion $[\mathfrak{g}, \mathfrak{g}_{i-1}] \subset \mathfrak{g}_i$ (de sorte qu'en particulier toute sous-algèbre $\mathfrak{g}_{i-1}$ est un idéal de l'algèbre $\mathfrak{g}$).

Démonstration. (Comparer avec la démonstration de la proposition 1.) Si $\mathfrak{g}^i \subset \mathfrak{g}_{i-1}$, alors $\mathfrak{g}^{i+1} = [\mathfrak{g}, \mathfrak{g}^i] \subset [\mathfrak{g}, \mathfrak{g}_{i-1}] \subset \mathfrak{g}_i$. Comme $\mathfrak{g}^1 = \mathfrak{g}_0$, ceci prouve par récurrence que $\mathfrak{g}^i \subset \mathfrak{g}_{i-1}$ pour tout $i = 1, \ldots, n$. En particulier, $\mathfrak{g}^{n+1} \subset \mathfrak{g}_n = 0$ et, par suite, $\mathfrak{g}^{n+1} = 0$.

Réciproquement, s'il existe un $k \geqslant 1$ tel que $\mathfrak{g}^k = 0$, alors (pourvu que $\mathfrak{g}^{k-1} \neq 0$) les inclusions sont toutes strictes dans la suite

$$\mathfrak{g} = \mathfrak{g}^1 \supset \mathfrak{g}^2 \supset \ldots \supset \mathfrak{g}^k = 0$$

et, par conséquent, cette suite peut être incluse dans le drapeau de sous-espaces

$$\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \ldots \supset \mathfrak{g}_n = 0.$$

20—01642

Si maintenant a est le plus grand indice pour lequel $\mathfrak{g}_i \subset \mathfrak{g}^a$, alors $\mathfrak{g}^{a+1} \subset \mathfrak{g}_{i+1}$ et, par suite,

$$[\mathfrak{g}, \mathfrak{g}_i] \subset [\mathfrak{g}, \mathfrak{g}^a] = \mathfrak{g}^{a+1} \subset \mathfrak{g}_{i+1}. \quad \square$$

On démontre comme pour les algèbres résolubles que *toute sous-algèbre et toute algèbre quotient d'une algèbre nilpotente sont des algèbres nilpotentes.* Mais cette proposition n'est généralement pas valable pour les extensions. On peut tout au plus affirmer qu'*une algèbre de Lie* $\mathfrak{g}$ *est nilpotente s'il en est de même de son algèbre quotient* $\mathfrak{g}/\mathfrak{h}$ *par un idéal* $\mathfrak{h}$ *contenu dans son centre* $\mathfrak{z}$. En effet, si $(\mathfrak{g}/\mathfrak{h})^k = 0$, alors $\mathfrak{g}^k \subset \mathfrak{h} \subset \mathfrak{z}$ et, par suite, $\mathfrak{g}^{k+1} = [\mathfrak{g}, \mathfrak{g}^k] \subset [\mathfrak{g}, \mathfrak{z}] = 0$. $\square$

Si $\mathfrak{a}$ est un idéal d'une algèbre de Lie $\mathfrak{g}$, ses idéaux $\mathfrak{a}^k$ sont de toute évidence des idéaux de $\mathfrak{g}$. De façon plus générale, étant donnés deux idéaux $\mathfrak{a}$ et $\mathfrak{b}$, leurs sous-espaces

$$\mathfrak{c}_0 = \mathfrak{d}_0, \ \mathfrak{c}_1 = [\mathfrak{c}_0, \mathfrak{d}_1], \ \ldots, \ \mathfrak{c}_i = [\mathfrak{c}_{i-1}, \mathfrak{d}_i],$$

où $\mathfrak{d}_0, \mathfrak{d}_1, \ldots, \mathfrak{d}_i, \ldots$ sont des idéaux confondus chacun soit avec $\mathfrak{a}$, soit avec $\mathfrak{b}$, seront aussi des idéaux de l'algèbre $\mathfrak{g}$. Ceci étant, il est aisé de voir que *pour tout* $k \geqslant 0$, *on a l'inclusion*

$$\mathfrak{c}_k \subset \mathfrak{a}^l,$$

où l *est le nombre d'indices* $i \leqslant k$ *tels que* $\mathfrak{d}_i = \mathfrak{a}$. En effet, pour $k = 0$ cette inclusion est évidente (on convient que $\mathfrak{a}^0 = \mathfrak{g}$), et si elle est vraie pour un certain k, alors $\mathfrak{c}_{k+1} = [\mathfrak{c}_k, \mathfrak{d}_k] \subset [\mathfrak{a}^l, \mathfrak{d}_k]$ et, par suite, $\mathfrak{c}_{k+1} \subset \mathfrak{a}^l$ si $\mathfrak{d}_k = \mathfrak{b}$ et $\mathfrak{c}_{k+1} \subset \mathfrak{a}^{l+1}$ si $\mathfrak{d}_{k+1} = \mathfrak{a}$. $\square$

Par symétrie on a visiblement l'inclusion $\mathfrak{c}_k \subset \mathfrak{b}^m$, où m est le nombre d'indices $i \leqslant k$ pour lesquels $\mathfrak{d}_i = \mathfrak{b}$. Mais il est clair que, ou bien l, ou bien m est $\geqslant p = [k/2]$. Donc, soit $\mathfrak{c}_k \subset \mathfrak{a}^p$, soit $\mathfrak{c}_k \subset \mathfrak{b}^p$, c'est-à-dire que

$$\mathfrak{c}_k \subset \mathfrak{a}^p \cup \mathfrak{b}^p.$$

Par ailleurs, il est évident que l'idéal $(\mathfrak{a} + \mathfrak{b})^k$ est somme des idéaux de la forme $\mathfrak{c}_k$ correspondant à toutes les suites $\mathfrak{d}_0, \mathfrak{d}_1, \ldots, \mathfrak{d}_k$ des idéaux $\mathfrak{a}$ et $\mathfrak{b}$. Donc, cet idéal satisfait l'inclusion

$$(\mathfrak{a} + \mathfrak{b})^k \subset \mathfrak{a}^p \cup \mathfrak{b}^p.$$

De là il s'ensuit, en particulier, que de même que pour les idéaux résolubles, *la somme* $\mathfrak{a} + \mathfrak{b}$ *des idéaux nilpotents* $\mathfrak{a}$ *et* $\mathfrak{b}$ *est un idéal nilpotent.* Donc, *dans toute algèbre de Lie* $\mathfrak{g}$ *il existe un plus grand idéal nilpotent* $\mathfrak{n}$, contenant tous les autres idéaux nilpotents de $\mathfrak{g}$. Cet idéal s'appelle *le nilradical* de l'algèbre de Lie $\mathfrak{g}$.

A noter que contrairement au cas du radical, l'algèbre quotient $\mathfrak{g}/\mathfrak{n}$ d'une algèbre de Lie $\mathfrak{g}$ par son nilradical $\mathfrak{n}$ peut très bien avoir un nilradical non nul.

Pour les sous-algèbres de l'algèbre des commutateurs $[\mathcal{A}]$ d'une algèbre associative $\mathcal{A}$ (généralement de dimension infinie) et, en particulier, pour les *algèbres de Lie linéaires* (les sous-algèbres de l'algèbre des commutateurs $[\text{End } \mathcal{V}]$ des opérateurs linéaires définis sur un espace vectoriel $\mathcal{V}$), on peut exhiber une condition suffisante de nilpotence qui est très utile.

On rappelle qu'un opérateur linéaire défini sur un espace vectoriel $\mathcal{V}$ (ou, de manière plus générale, un élément d'une algèbre associative $\mathcal{A}$) est *nilpotent* si l'une quelconque de ses puissances est nulle (cf. II, 15). De façon analogue, l'ensemble des opérateurs linéaires (ou des éléments d'une algèbre associative $\mathcal{A}$) est *nilpotent* s'il existe un $k \geqslant 1$ tel que le produit de k quelconques de ses éléments est nul.

Nous appliquerons cette dernière notion aux sous-ensembles qui sont sous-espaces (et notamment sous-algèbres) de l'algèbre de Lie $[\mathcal{A}]$, donc, pour éviter toute confusion terminologique, on appellera les sous-espaces nilpotents en ce sens *sous-espaces* (*sous-algèbres*) *associativement nilpotents* (*nilpotentes*). Par ailleurs, conformément à la terminologie en usage en théorie des algèbres associatives, les sous-algèbres $\mathfrak{g} \subset [\mathcal{A}]$ composées d'éléments nilpotents seront appelées *nilsous-algèbres de Lie*. Il est évident que toute sous-algèbre de Lie associativement nilpotente est une nilsous-algèbre. Fait remarquable, la réciproque est vraie:

Proposition 3. *Toute nilsous-algèbre de Lie $\mathfrak{g}$ de dimension finie est associativement nilpotente.*

Prouvons une proposition plus générale.

Proposition 3* (théorème de Jacobson). *Soient $\mathcal{A}$ une algèbre associative (éventuellement de dimension infinie) et $\mathfrak{G}$ son sous-espace de dimension finie engendré par un sous-ensemble $\mathfrak{g}$ stable pour la commutation. Si les éléments $a \in \mathfrak{g}$ sont nilpotents, le sous-espace $\mathfrak{G}$ est associativement nilpotent.*

D é m o n s t r a t i o n. Il est aisé de voir que $\mathfrak{g}$ contient des sous-ensembles associativement nilpotents (par exemple, le sous-ensemble nul). L'ensemble $\mathfrak{G}$ étant par hypothèse de rang fini, il contient des sous-ensembles $\mathfrak{h}$ maximaux associativement nilpotents. Le sous-espace engendré par un sous-ensemble associativement nilpotent étant visiblement associativement nilpotent, il suffit de montrer que $\mathfrak{h} = \mathfrak{g}$. Supposons que $\mathfrak{h} \neq \mathfrak{g}$ et montrons que ceci est absurde.

Soit $\mathcal{P}$ l'enveloppe linéaire d'un sous-ensemble $\mathfrak{h}$ maximal associativement nilpotent, et soit p un nombre tel que le produit de p éléments de $\mathcal{P}$ soit nul. Alors *le résultat de la commutation successive d'un élément $a \in \mathfrak{g}$ à $2p - 1$ éléments quelconques de $\mathcal{P}$ est nul.* En effet, la commutation successive de a à des éléments $b_1, \ldots$

. . ., b_{2p-1} de $\mathcal{P}$ nous donne, après développement des crochets de Lie, une somme algébrique de produits de la forme bac, où b est le produit de quelques éléments $b_1, \ldots, b_{2p-1}$ et c, le produit des autres éléments. Mais, il est clair que ou bien b, ou bien c contient pas moins de p facteurs $b_1, \ldots, b_{2-p1}$, donc ce produit est nul. Par conséquent, tous les éléments de la forme bac sont nuls, donc leur somme aussi.

Il est évident maintenant que *pour tout élément* $a \in \mathfrak{g}$ *il existe un nombre* $s \geqslant 0$ *tel que le résultat de la commutation successive de* a *à* s *éléments de* $\mathfrak{h}$ *appartient à* $\mathcal{P}$. Ceci étant, $s \leqslant 2p - 1$.

De là il s'ensuit que *si* $\mathfrak{h} \neq \mathfrak{g}$, *alors il existe dans* $\mathfrak{g}$ *un élément* $a_0 \notin \mathfrak{h}$ *tel que* $[a_0, b] \in \mathcal{P}$ *pour tout* $b \in \mathcal{P}$. Pour le prouver il suffit de prendre un élément quelconque $a \in \mathfrak{g} \setminus \mathfrak{h}$ et de lui appliquer la proposition établie ci-dessus, en remarquant que si s est minimal, il existe dans $\mathfrak{h}$ des éléments $b_1, \ldots, b_{s-1}$ tels que le résultat a_0 (appartenant à $\mathfrak{g}$) de la commutation successive de a à $b_1, \ldots, b_{s-1}$ n'appartient pas à $\mathfrak{h}$, mais est tel que $[a_0, b] \in \mathcal{P}$ pour tout $b \in \mathfrak{h}$, donc et pour tout $b \in \mathcal{P}$.

Appelons *monôme* un produit d'éléments égaux à a_0 ou appartenant à $\mathcal{P}$. Soit r le nombre de facteurs de $\mathcal{P}$ d'un monôme a. Il s'avère que *si* $r \geqslant p$, *alors* $a = 0$. En effet, supposons que le monôme a contienne un facteur de la forme ba_0, où $b \in \mathcal{P}$. Comme $[a_0, b] \in \in \mathcal{P}$, en remplaçant ba_0 par $a_0b - [a_0, b]$, on représente a par la somme de deux monômes avec le même r, le premier renfermant un facteur a_0 de moins, alors que dans le second ce facteur a_0 se déplacera d'un rang à gauche. En effectuant cette procédure le nombre de fois nécessaire, on mettra a sous la forme d'une somme de monômes dans lesquels soit les facteurs a_0 auront disparu, soit ils seront tous rassemblés à gauche. Mais chacun de ces monômes s'obtient en faisant le produit d'un élément de l'algèbre $\mathcal{A}$ par un élément qui est le produit de $r \geqslant p$ éléments de $\mathcal{P}$, donc est nul. Par conséquent, il en est de même de a.

Comme $a_0 \in \mathfrak{g}$, il existe par hypothèse une puissance k_0 telle que $a_0^{k_0} = 0$. Donc, le monôme a sera non nul si seulement il contient l'élément a_0 à des puissances $< k_0$. Ces puissances alternent avec les facteurs de $\mathcal{P}$ et pour cette raison sont en nombre inférieur à r. Comme, d'après ce qui précède, pour $a \neq 0$ on a nécessairement $r < p$, le nombre total des facteurs de tout monôme $a \neq 0$ est inférieur à

$$r - 1 + r(k_0 - 1) \leqslant p - 1 + p(k_0 - 1) = pk_0 - 1.$$

Ceci prouve que *si les facteurs de* a *sont en nombre supérieur á* pk_0, *alors* $a = 0$.

En particulier, cela signifie que le produit de pk_0 éléments quelconques de l'ensemble $\bar{\mathfrak{h}} = \{\mathfrak{h}, a_0\}$ est nul, c'est-à-dire que cet ensemble est associativement nilpotent. Comme ceci contredit le

fait que l'ensemble $\mathfrak{h}$ est maximal, la proposition 3* est entièrement prouvée. □

Corollaire 1. *Toute nilsous-algèbre de Lie* $\mathfrak{g} \subset [\mathscr{A}]$ *de dimension finie est nilpotente.*

Démonstration. Chaque élément de l'idéal $\mathfrak{g}^k$ est une combinaison linéaire de crochets de Lie itérés de la forme

$$[x_1, [x_2, \ldots [x_{k-1}, x_k] \ldots]], \qquad x_1, x_2, \ldots, x_k \in \mathfrak{g},$$

et, par suite (en tant qu'élément de l'algèbre $\mathscr{A}$), est somme algébrique de tous les produits possibles de la forme $x_{i_1} \ldots x_{i_k}$. Donc, si une algèbre de Lie $\mathfrak{g}$ est associativement nilpotente, alors $\mathfrak{g}^k = 0$ pour k assez grand, c'est-à-dire que cette algèbre est nilpotente. Pour achever la démonstration, il nous reste donc à utiliser la proposition 3. □

En général, cette condition suffisante de nilpotence de sous-algèbres d'algèbres de Lie de commutateurs (et, en particulier, d'algèbres de Lie linéaires) n'est pas nécessaire. Il est aisé de voir par exemple que l'ensemble de toutes les matrices de la forme $\lambda E + A$, où $A = (a_{ij})$ est une matrice trigonale supérieure stricte (c'est-à-dire telle que $a_{ij} = 0$ pour $i \leqslant j$), est sous-algèbre nilpotente de l'algèbre de Lie $[\mathbb{R}(n)]$ et la matrice $\lambda E + A$ n'est nilpotente que pour $\lambda = 0$.

Revenons maintenant aux algèbres de Lie (de dimension finie).

Pour toute algèbre de Lie $\mathfrak{g}$ est défini un homomorphisme ad dans l'algèbre des commutateurs de l'algèbre des opérateurs linéaires $\mathfrak{g} \to \mathfrak{g}$ (cf. leçon 3). Par définition, pour tout élément $a \in \mathfrak{g}$, l'opérateur linéaire ad a: $\mathfrak{g} \to \mathfrak{g}$ envoie $x \in \mathfrak{g}$ dans $[a, x]$. Donc, en particulier, tout crochet de Lie itéré $[x_1, [x_2, \ldots, [x_{k-1}, x_k] \ldots]]$ n'est autre que l'image de l'élément x_k par l'opérateur linéaire ad $x_1 \circ$ ad $x_2 \circ \ldots \circ$ ad x_{k-1}. Cela signifie qu'*une algèbre de Lie* $\mathfrak{g}$ *est nilpotente si et seulement si l'algèbre de Lie linéaire* ad $\mathfrak{g}$ *est associativement nilpotente.* En vertu de la proposition 3, on a ainsi prouvé le

Corollaire 2. *Une algèbre de Lie* $\mathfrak{g}$ *est nilpotente si et seulement si l'opérateur* ad a *est nilpotent pour tout élément* $a \in \mathfrak{g}$. □

Ce corollaire est connu sous le nom de *théorème d'Engel.* Du reste, on désigne très souvent le corollaire 1 et même la proposition 3 par le théorème d'Engel.

Les critères de nilpotence d'un opérateur linéaire revêtent une signification particulière avec le théorème d'Engel. On se propose d'en démontrer deux qui sont élémentaires.

On rappelle que la *trace* d'un opérateur linéaire A défini sur un espace vectoriel $\mathcal{V}$ est la somme des éléments diagonaux de sa matrice dans une base quelconque de $\mathcal{V}$. La trace est définie intrinsèquement (ne dépend pas de la base) et se note $\operatorname{Tr} A$.

Propriétés de la trace.

1° Le nombre $\operatorname{Tr} A$ dépend linéairement de l'opérateur A (est une fonctionnelle linéaire sur l'espace End $\mathcal{V}$ des opérateurs linéaires $\mathcal{V} \to \mathcal{V}$).

2° Quels que soient les opérateurs A et B, on a

$$\operatorname{Tr} AB = \operatorname{Tr} BA.$$

3° La trace est égale à la somme des valeurs propres de l'opérateur A (répétées un nombre de fois égal à leur multiplicité):

$$\operatorname{Tr} A = \lambda_1 + \ldots + \lambda_n.$$

La propriété 1° est évidente. La propriété 2° se démontre par un calcul direct. Pour établir la propriété 3°, le plus simple est de considérer la forme normale de Jordan de l'opérateur A. (A noter que dans la propriété 3° on passe à la clôture algébrique du corps $\mathbb{K}$, c'est-à-dire si $\mathbb{K} = \mathbb{R}$, au corps $\mathbb{C}$, mais la trace de l'opérateur ne change pas si l'on élargit le corps de base, puisque la matrice ne change pas.)

Comme les valeurs propres d'un opérateur nilpotent sont nulles (cf. II, 15), la propriété 3° entraîne, en particulier, que *la trace de tout opérateur nilpotent est nulle.*

Cette condition nécessaire de nilpotence n'est évidemment pas suffisante. Cependant, comme toute puissance d'un opérateur nilpotent est aussi un opérateur nilpotent, il s'ensuit que *si A est un opérateur nilpotent, alors* $\operatorname{Tr} A^k = 0$ *pour tout k.*

Il s'avère que cette condition est nécessaire et suffisante, c'est-à-dire que *si* $\operatorname{Tr} A^k = 0$ *pour tout k, alors l'opérateur A est nilpotent.* En effet, étant donné que les valeurs propres de l'opérateur A^k sont les puissances $\lambda_1^k, \ldots, \lambda_n^k$ des valeurs propres de A, la trace $\operatorname{Tr} A^k$ de A^k vérifie la formule

$$\operatorname{Tr} A^k = \lambda_1^k + \ldots + \lambda_n^k$$

qui exprime que $\operatorname{Tr} A^k$ est la somme des puissances k-ièmes des valeurs propres du polynôme caractéristique de A. Mais, de la théorie des polynômes symétriques on sait que les coefficients de tout polynôme s'expriment polynomialement en fonction de la somme des puissances de ses racines (*formules de Waring*) et sont nuls si toutes ces sommes le sont. En appliquant ceci au polynôme caractéristique $f_A(\lambda)$ de l'opérateur A, on trouve que si $\operatorname{Tr} A^k = 0$ pour tout k, alors $f_A(\lambda) = \lambda^n$. Donc, $A^n = 0$ en vertu du théorème de Cayley-Hamilton (cf. II, 16). □

Une autre condition suffisante plus spéciale de nilpotence se rapporte aux opérateurs de la forme

$$A = [B_1, C_1] + \ldots + [B_s, C_s] \tag{1}$$

qui sont des sommes de crochets de Lie. Il s'avère que *si un opérateur A de la forme* (1) *est permutable à tout opérateur $B_1, \ldots, B_s$ (c'est-à-dire que $[A, B_1] = 0, \ldots, [A, B_s] = 0$), alors cet opérateur est nilpotent.* En effet, pour tout k l'opérateur A^k est la somme des crochets de Lie

$$\begin{aligned} A^k &= A^{k-1}([B_1, C_1] + \ldots + [B_s, C_s]) = \\ &= A^{k-1}(B_1C_1 - C_1B_1 + \ldots + B_sC_s - C_sB_s) = \\ &= (B_1A^{k-1}C_1 - A^{k-1}C_1B_1) + \ldots + (B_sA^{k-1}C_s - A^{k-1}C_sB_s) = \\ &= [B_1, A^{k-1}C_1] + \ldots + [B_s, A^{k-1}C_s], \end{aligned}$$

donc $\operatorname{Tr} A^k = 0$ (cf. propriétés 1 et 2 de la trace). Donc, l'opérateur A est nilpotent. □

De ce critère il résulte que *l'intersection $\mathfrak{z} \cap \mathfrak{g}^2$ du centre d'une algèbre de Lie linéaire $\mathfrak{g}$ et de l'idéal $\mathfrak{g}^2$ se compose d'opérateurs nilpotents.* En effet, tout opérateur A de $\mathfrak{g}^2$ est de la forme (1), où $B_1, C_1, \ldots, B_s, C_s \in \mathfrak{g}$ et si $A \in \mathfrak{z}$, alors $[A, B_1] = 0, \ldots$ $\ldots, [A, B_s] = 0$. □

On démontre de façon analogue que *pour tout idéal abélien $\mathfrak{a}$ d'une algèbre de Lie linéaire $\mathfrak{g}$, l'idéal $[\mathfrak{a}, \mathfrak{g}]$ se compose d'opérateurs nilpotents.*

On obtient des résultats plus précis si l'on admet que l'algèbre de Lie linéaire $\mathfrak{g}$ est *irréductible*, c'est-à-dire qu'il n'existe pas dans $\mathcal{V}$ de sous-espaces non triviaux invariants par les opérateurs de $\mathfrak{g}$.

Pour tout sous-ensemble $\mathfrak{a}$ d'une algèbre de Lie linéaire $\mathfrak{g}$ opérant dans un espace $\mathcal{V}$, on désignera par $\mathfrak{a}\mathcal{V}$ l'enveloppe linéaire des vecteurs de la forme Ax, où $A \in \mathfrak{a}$ et $x \in \mathcal{V}$.

Il est immédiat de voir que *si $\mathfrak{a}$ est un idéal, le sous-espace $\mathfrak{a}\mathcal{V}$ est invariant par les opérateurs de $\mathfrak{g}$.* En effet, si $A \in \mathfrak{a}$, $B \in \mathfrak{g}$ et $x \in \mathcal{V}$, alors

$$B(Ax) = [B, A]x + A(Bx) \in \mathfrak{a}\mathcal{V},$$

car $[B, A] \in \mathfrak{a}$. □

Il s'ensuit de là que *si une algèbre de Lie linéaire $\mathfrak{g}$ est irréductible, alors chacun de ses idéaux $\mathfrak{a}$ non nuls contient un opérateur nilpotent.* En effet, si les opérateurs de $\mathfrak{a}$ sont tous nilpotents, c'est-à-dire si $\mathfrak{a}$ est une algèbre nilpotente, alors la proposition 3 nous dit que l'idéal $\mathfrak{a}$ est associativement nilpotent. Soit m le plus petit nombre tel que le produit de m éléments quelconques de $\mathfrak{a}$ soit nul. Alors,

dans la série de sous-espaces

$$\mathcal{P}_1 = \mathfrak{a}\mathcal{V}, \quad \mathcal{P}_2 = \mathfrak{a}\mathcal{P}_1, \; \ldots, \quad \mathcal{P}_{m-1} = \mathfrak{a}\mathcal{P}_{m-2}, \quad \mathcal{P}_m = \mathfrak{a}\mathcal{P}_{m-1}$$

le sous-espace $\mathcal{P}_m$ est nul et $\mathcal{P}_{m-1} \neq 0$. Ceci étant, en vertu de l'assertion précédente (appliquée $m - 1$ fois), le sous-espace $\mathcal{P}_{m-1}$ est invariant par tous les opérateurs de $\mathfrak{g}$. Donc, $\mathcal{P}_{m-1} = \mathcal{V}$ en vertu de l'irréductibilité et, par suite, $\mathcal{P} = \mathfrak{a}\mathcal{V} = \mathfrak{a}\mathcal{P}_{m-1} = \mathcal{P}_m = 0$, ce qui n'est possible que pour $\mathfrak{a} = 0$. □

En appliquant cette assertion à l'idéal $\mathfrak{z} \cap \mathfrak{g}^2$, on déduit aussitôt que *pour toute algèbre de Lie irréductible linéaire* $\mathfrak{g}$ *on a*

$$\mathfrak{z} \cap \mathfrak{g}^2 = 0.$$

On voit par ailleurs que *tout idéal abélien* $\mathfrak{a}$ *d'une algèbre de Lie* $\mathfrak{g}$ *irréductible linéaire est contenu dans le centre* $\mathfrak{z}$ *de* $\mathfrak{g}$. En effet, l'idéal $[\mathfrak{a}, \mathfrak{g}]$ est nul, car ne contenant que des opérateurs nilpotents. Or l'égalité $[\mathfrak{a}, \mathfrak{g}] = 0$ signifie très exactement que $\mathfrak{a} \subset \mathfrak{z}$. □

Définition 5. On dit qu'une algèbre de Lie $\mathfrak{g}$ est *réductive* si ses idéaux abéliens appartiennent à son centre $\mathfrak{z}$ et $\mathfrak{z} \cap \mathfrak{g}^2 = 0$.

Nous avons ainsi montré que *toute algèbre irréductible linéaire est réductive.*

Soit $\mathfrak{g}$ une algèbre de Lie réductive. Comme $\mathfrak{z} \cap \mathfrak{g}^2 = 0$, on peut élargir $\mathfrak{g}^2$ à un sous-espace $\mathfrak{m}$, tel que $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{m}$. De plus, le sous-espace $\mathfrak{m}$ est un idéal, car $[\mathfrak{g}, \mathfrak{m}] \subset [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{m}$. Par ailleurs, il est immédiat de voir que tout idéal $\mathfrak{a}$ de $\mathfrak{m}$ est idéal dans $\mathfrak{g}$ (puisque $\mathfrak{g} = \mathfrak{z} + \mathfrak{m}$, alors $[\mathfrak{g}, \mathfrak{a}] \subset [\mathfrak{g}, \mathfrak{a}] + [\mathfrak{m}, \mathfrak{a}] = [\mathfrak{m}, \mathfrak{a}] \subset \mathfrak{a}$). Donc, si l'idéal $\mathfrak{a} \subset \mathfrak{m}$ est abélien, alors $\mathfrak{a} \subset \mathfrak{z}$ et, par suite, $\mathfrak{a} = 0$. Donc, l'idéal $\mathfrak{m}$ ne possède pas d'idéaux abéliens non nuls et de ce fait est semi-simple.

Réciproquement, si l'algèbre de Lie $\mathfrak{g}$ est de la forme $\mathfrak{z} \oplus \mathfrak{m}$, où $\mathfrak{m}$ est un idéal semi-simple, alors $\mathfrak{g}^2 = \mathfrak{m}^2 \subset \mathfrak{m}$ et $\mathfrak{z} \cap \mathfrak{g}^2 = 0$. Par ailleurs, l'épimorphisme canonique $\mathfrak{g} \to \mathfrak{g}/\mathfrak{z}$ envoie tout idéal abélien $\mathfrak{a}$ de l'algèbre $\mathfrak{g}$ dans un idéal abélien de l'algèbre semi-simple $\mathfrak{g}/\mathfrak{z} \approx \mathfrak{m}$, c'est-à-dire dans 0. Donc, $\mathfrak{a} \subset \mathfrak{z}$ et, par suite, l'algèbre de Lie $\mathfrak{g}$ est réductive.

Ce qui prouve qu'*une algèbre de Lie* $\mathfrak{g}$ *est réductive si et seulement si elle est somme directe de son centre et d'un idéal semi-simple* $\mathfrak{m}$:

$$\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{m}.$$

(A noter qu'en fait $\mathfrak{m} = \mathfrak{g}^2$. En effet, de $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{m}$ il résulte que $\mathfrak{g}^2 = \mathfrak{m}^2$, et, comme nous le prouverons dans la leçon suivante, pour toute algèbre de Lie semi-simple $\mathfrak{m}$, on a $\mathfrak{m}^2 = \mathfrak{m}$.)

Il est immédiat de voir maintenant que *le radical* $\mathfrak{r}$ *d'une algèbre de Lie* $\mathfrak{g}$ *réductive est confondu avec son centre* $\mathfrak{z}$. En effet, le radical $\mathfrak{r}$ se transforme en 0 par l'épimorphisme canonique $\mathfrak{g} \to \mathfrak{g}/\mathfrak{z}$ (car l'algè-

bre $\mathfrak{g}/\mathfrak{z}$ est isomorphe à l'idéal $\mathfrak{m}$, donc est semi-simple). Par conséquent, $\mathfrak{r} \subset \mathfrak{z}$ et $\mathfrak{r} = \mathfrak{z}$. □

Donc, *une algèbre de Lie réductive est résoluble si et seulement si elle est abélienne.*

En particulier donc, *toute algèbre de Lie résoluble irréductible linéaire est abélienne.*

Pour appliquer cette assertion à des algèbres pas nécessairement irréductibles, nous prouverons préalablement un lemme général d'algèbre linéaire relatif à un opérateur linéaire A agissant sur un espace vectoriel $\mathcal{V}$ et laissant invariant un sous-espace $\mathcal{P}$ de $\mathcal{V}$. On sait qu'un tel opérateur induit un opérateur $A_1: \mathcal{P} \to \mathcal{P}$ et un opérateur $A_2: \mathcal{V}/\mathcal{P} \to \mathcal{V}/\mathcal{P}$.

Lemme 1. *Si les opérateurs A_1 et A_2 sont nilpotents, il en est de même de A.*

D é m o n s t r a t i o n. Soient $A_1^{k_1} = 0$ et $A_2^{k_2} = 0$. Alors A^{k_2} envoie $\mathcal{V}$ dans $\mathcal{P}$ et A^{k_1}, $\mathcal{P}$ en 0. Donc $A^{k_1+k_2} = A^{k_1}A^{k_2}$ envoie $\mathcal{V}$ en 0, c'est-à-dire que $A^{k_1+k_2} = 0$. □

Il est immédiat de voir maintenant que *si $\mathfrak{g}$ est une algèbre de Lie résoluble linéaire et A un opérateur nilpotent de $\mathfrak{g}$, alors quel que soit l'opérateur B de $\mathfrak{g}$ l'opérateur AB est nilpotent.* En effet, raisonnons par récurrence sur la dimension n de l'espace vectoriel $\mathcal{V}$ dans lequel opère l'algèbre de Lie $\mathfrak{g}$, en tenant compte du fait que cette assertion est triviale pour $n = 1$. Si $\mathcal{V}$ ne contient pas de sous-espaces non triviaux invariants par les opérateurs de $\mathfrak{g}$, c'est-à-dire si l'algèbre $\mathfrak{g}$ est irréductible, alors elle est abélienne d'après ce qui a été démontré ci-dessus et, par suite, $AB = BA$. Donc, $(AB)^k = A^kB^k$ pour tout $k \geqslant 0$ et, par suite, $(AB)^k = 0$ lorsque $A^k = 0$. Si dans $\mathcal{V}$ il existe un sous-espace invariant non trivial $\mathcal{P}$, alors en considérant les restrictions des opérateurs de $\mathfrak{g}$ à $\mathcal{P}$, on obtient dans $\mathcal{P}$ une algèbre d'opérateurs qui est homomorphe à $\mathfrak{g}$ et donc est résoluble. Donc, d'après l'hypothèse de récurrence, la restriction de l'opérateur AB à $\mathcal{P}$ est nilpotente. De façon analogue, en passant à l'espace quotient $\mathcal{V}/\mathcal{P}$, on trouve que l'opérateur AB induit un opérateur nilpotent sur ce sous-espace. Donc il est lui-même nilpotent d'après le lemme 1. □

A noter qu'en général $AB \notin \mathfrak{g}$.

On démontre de façon analogue que *pour toute algèbre de Lie linéaire $\mathfrak{g}$ de radical* $\mathfrak{r}$, *l'idéal* $[\mathfrak{g}, \mathfrak{r}]$ *se compose d'opérateurs nilpotents.* En effet, si l'algèbre $\mathfrak{g}$ est irréductible, on sait que $\mathfrak{r} = \mathfrak{z}$, et, par suite, $[\mathfrak{g}, \mathfrak{r}] = 0$. Donc, dans ce cas cette assertion est automatiquement vraie. Dans le cas général, nous procéderons encore par récurrence sur $\dim \mathcal{V}$. Soit $\mathcal{P}$ le sous-espace de $\mathcal{V}$ invariant

par les opérateurs de $\mathfrak{g}$. En considérant alors les restrictions des opérateurs de $\mathfrak{g}$ à $\mathcal{P}$, on obtient dans $\mathcal{P}$ une algèbre d'opérateurs $\mathfrak{g}'$ isomorphe à l'algèbre quotient de $\mathfrak{g}$ par un idéal (composé d'opérateurs nuls sur $\mathcal{P}$). D'après l'hypothèse de récurrence, l'idéal $[\mathfrak{g}', \mathfrak{r}']$, où $\mathfrak{r}'$ est le radical de $\mathfrak{g}'$, se compose d'opérateurs nilpotents. Comme l'épimorphisme $\mathfrak{g} \to \mathfrak{g}'$ envoie l'idéal $[\mathfrak{g}, \mathfrak{r}]$ dans $[\mathfrak{g}', \mathfrak{r}']$, ceci prouve que tout opérateur A de $[\mathfrak{g}, \mathfrak{r}]$ induit dans $\mathcal{P}$ un opérateur nilpotent. On démontre de façon analogue que l'opérateur A induit un opérateur nilpotent également dans l'espace quotient $\mathcal{V}/\mathcal{P}$. Donc, l'opérateur A est nilpotent en vertu du lemme 1. □

De là on déduit, en vertu du théorème d'Engel, que pour toute algèbre de Lie linéaire $\mathfrak{g}$, l'idéal $[\mathfrak{g}, \mathfrak{r}]$ est nulpotent. Du reste, il est aisé de voir que cette assertion est valable pour des algèbres de Lie arbitraires.

Proposition 4. *L'idéal* $[\mathfrak{g}, \mathfrak{r}]$ *de toute algèbre de Lie* $\mathfrak{g}$ *est nilpotent.*

D é m o n s t r a t i o n. Soit $\mathfrak{g}' = \operatorname{ad} \mathfrak{g}$, et soit $\mathfrak{r}'$ le radical de l'algèbre de Lie $\mathfrak{g}'$. D'après ce qui précède, $[\mathfrak{g}', \mathfrak{r}']$ est nilpotent. Par ailleurs, l'homomorphisme $\operatorname{ad}: \mathfrak{g} \to \mathfrak{g}'$ transforme l'idéal $[\mathfrak{g}, \mathfrak{r}]$ justement en l'idéal $[\mathfrak{g}', \mathfrak{r}']$ et le noyau de cet homomorphisme est le centre de l'algèbre $\mathfrak{g}$. Donc, l'algèbre quotient de l'idéal $[\mathfrak{g}, \mathfrak{r}]$ par un idéal du centre est une algèbre nilpotente. Donc, l'idéal $[\mathfrak{g}, \mathfrak{r}]$ est lui-même nilpotent. □

Corollaire. *Une algèbre de Lie* $\mathfrak{g}$ *est résoluble si et seulement si l'idéal* $\mathfrak{g}^2$ *est nilpotent.*

D é m o n s t r a t i o n. Si l'algèbre $\mathfrak{g}$ est résoluble, c'est-à-dire si $\mathfrak{g} = \mathfrak{r}$, alors l'idéal $\mathfrak{g}^2 = [\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{r}]$ est nilpotent. Réciproquement, si l'idéal $\mathfrak{g}^2$ est nilpotent (et, par suite, résoluble), l'algèbre $\mathfrak{g}$ est résoluble, puisque l'algèbre quotient $\mathfrak{g}/\mathfrak{g}^2$ est abélienne. □

L'idéal $[\mathfrak{g}, \mathfrak{r}]$ s'appelle généralement *radical nilpotent* de l'algèbre de Lie $\mathfrak{g}$. Il est contenu dans le nilradical $\mathfrak{n}$ et en général il est distinct de $\mathfrak{n}$.

LEÇON 18

Fonctionnelle de trace.— Fonctionnelle de Killing.— Fonctionnelle de trace d'une représentation.— Décomposition de Jordan d'un opérateur linéaire.— Décomposition de Jordan d'un opérateur adjoint.— Théorème de Cartan sur les algèbres de Lie linéaires.— Démonstration du critère de Cartan de résolubilité d'une algèbre de Lie.— Algèbres de Lie linéaires à fonctionnelle de trace non dégénérée.— Algèbres de Lie semi-simples.— Critère de Cartan de semi-simplicité.— Opérateurs de Casimir.

Poursuivons l'étude des algèbres de Lie linéaires commencée dans la leçon précédente.

Des propriétés de la trace il s'ensuit immédiatement que la formule

$$t(A, B) = \mathrm{Tr}\, AB$$

définit sur toute algèbre de Lie linéaire $\mathfrak{g}$ une fonctionnelle t symétrique bilinéaire. Cette fonctionnelle sera appelée *fonctionnelle de trace* de l'algèbre de Lie $\mathfrak{g}$.

On dit qu'une fonctionnelle bilinéaire s sur une algèbre de Lie $\mathfrak{g}$ est *invariante* (point n'est besoin d'expliquer ici l'origine de cette dénomination) si

$$s([x, y], z) = s(x, [y, z])$$

quels que soient $x, y, z \in \mathfrak{g}$, c'est-à-dire si tous les opérateurs linéaires ad y, $y \in \mathfrak{g}$, sont antisymétriques relativement à s.

Etant donné que quels que soient les opérateurs A, B et C, la trace de l'opérateur $[A, B]C = ABC - BAC$ est égale à celle de l'opérateur $A[B, C] = ABC - ACB$, *la fonctionnelle de trace de toute algèbre de Lie linéaire* $\mathfrak{g}$ *est invariante.*

Soit $\mathfrak{r}$ le radical d'une algèbre de Lie linéaire $\mathfrak{g}$, et soient $A, B \in \mathfrak{g}$ et $C \in \mathfrak{r}$. L'opérateur $[B, C]$ est nilpotent, puisque $[B, C] \in [\mathfrak{g}, \mathfrak{r}]$. Comme $\mathfrak{r}$ est un idéal, le sous-espace $\mathfrak{a}$ de $\mathfrak{g}$ engendré par $\mathfrak{r}$ et A

est tel que $[\mathfrak{a}, \mathfrak{a}] \subset \mathfrak{r}$. Donc, $\mathfrak{a}$ est une sous-algèbre de $\mathfrak{g}$ qui, de plus, est résoluble. Comme $A \in \mathfrak{a}$, $[B, C] \in [\mathfrak{g}, \mathfrak{r}] \subset \mathfrak{r} \subset \mathfrak{a}$, l'algèbre de Lie $\mathfrak{a}$ est résoluble et l'opérateur $[B, C]$ nilpotent, on déduit que l'opérateur $A[B, C]$ est nilpotent aussi. Donc, sa trace est nulle: $t(A, [B, C]) = 0$. Mais $t([A, B], C) = 0$ à cause de l'invariance. Ceci prouve que *dans toute algèbre de Lie linéaire* $\mathfrak{g}$, *les idéaux* $\mathfrak{g}^2$ *et* $\mathfrak{r}$ *sont orthogonaux relativement à la fonctionnelle de trace* t.

Dans une écriture conventionnelle, mais suggestive

$$t(\mathfrak{g}^2, \mathfrak{r}) = 0.$$

Pour $\mathfrak{g} = \mathfrak{r}$ on en déduit, en particulier, que *dans une algèbre de Lie* $\mathfrak{g}$ *résoluble linéaire on a*

$$t(\mathfrak{g}^2, \mathfrak{g}) = 0,$$

c'est-à-dire que l'idéal $\mathfrak{g}^2$ est orthogonal à l'algèbre $\mathfrak{g}$ tout entière relativement à la fonctionnelle de trace t.

Pour établir des résultats identiques pour une algèbre de Lie $\mathfrak{g}$ quelconque (généralement pas linéaire), nous passons à l'algèbre de Lie linéaire $\mathfrak{g}' = \operatorname{ad} \mathfrak{g}$. Transportons la fonctionnelle t définie sur cette algèbre linéaire dans l'algèbre de Lie $\mathfrak{g}$ par l'homomorphisme ad. En d'autres termes, définissons sur $\mathfrak{g}$ une fonctionnelle symétrique bilinéaire $t_{\mathfrak{g}}$ par la formule

$$t_{\mathfrak{g}}(x, y) = t(\operatorname{ad} x, \operatorname{ad} y) = \operatorname{Tr}(\operatorname{ad} x \operatorname{ad} y).$$

Définition 1. La fonctionnelle $t_{\mathfrak{g}}$ s'appelle *fonctionnelle de Killing* de l'algèbre de Lie $\mathfrak{g}$.

Comme ad est un homomorphisme d'algèbres de Lie, *la fonctionnelle de Killing est invariante.*

La fonctionnelle de Killing se calcule sans peine pour toute algèbre de Lie bien définie.

Exemple 1. Trouver la fonctionnelle de Killing de l'algèbre de Lie $\mathfrak{gl}(n)$ des matrices d'ordre n. Une base de cette algèbre est composée des matrices unités E_{ij}.

Comme $E_{ij}E_{\alpha\beta} = \delta_{j\alpha}E_{i,\beta}$ alors

$$(1) \qquad [X, E_{\alpha\beta}] = \sum_{i=1}^{n} (x_{i\alpha}E_{i\beta} - x_{\beta i}E_{\alpha i}), \qquad \alpha, \beta = 1, \ldots, n,$$

pour toute matrice $X = \sum_{i, j=1}^{n} x_{ij}E_{ij}$ de $\mathfrak{gl}(n)$, de sorte que

$$(\operatorname{ad} X) E_{\alpha\beta} = \sum_{i=1}^{n} (x_{i\alpha}E_{i\beta} - x_{\beta i}E_{\alpha i})$$

dans l'algèbre $\mathfrak{gl}(n)$.

Donc,

$$(\operatorname{ad} X \circ \operatorname{ad} Y) E_{\alpha\beta} = \sum_{i,j=1}^{n} (x_{i\alpha} y_{ji} E_{j\beta} + x_{\beta i} y_{ij} E_{\alpha j}) - \\ - \sum_{i,j=1}^{n} (x_{i\alpha} y_{\beta j} + x_{\beta j} y_{i\alpha}) E_{ij},$$

et, par suite,

$$\operatorname{Tr}(\operatorname{ad} X \circ \operatorname{ad} Y) = n \sum_{i,j=1}^{n} (x_{ij} y_{ji} + x_{ji} y_{ij}) - 2 \sum_{i,j=1}^{n} x_{ii} y_{jj} = \\ = 2n \operatorname{Tr}(XY) - 2 \operatorname{Tr} X \cdot \operatorname{Tr} Y.$$

Ceci prouve que *la fonctionnelle de Killing de l'algèbre de Lie* $\mathfrak{gl}(n)$ *s'exprime au moyen de la formule*

$$t_{\mathfrak{gl}(n)}(X, Y) = 2n \operatorname{Tr}(XY) - 2 \operatorname{Tr} X \cdot \operatorname{Tr} Y,$$

ou

$$t_{\mathfrak{gl}(n)}(X, Y) = 2nt(X, Y) - 2t(X, E) \cdot t(Y, E).$$

Signalons que *cette fonctionnelle est dégénérée*, c'est-à-dire que $\mathfrak{gl}(n)^{\perp} \neq 0$. En effet, il est clair que pour toute matrice scalaire aE, on a l'identité $t_{\mathfrak{gl}(n)}(X, aE) = 0$, $X \in \mathfrak{gl}(n)$, qui exprime que $aE \in \mathfrak{gl}(n)^{\perp}$.

Exemple 2. Une base de l'algèbre de Lie $\mathfrak{sl}(n)$ des matrices d'ordre n de trace nulle, est formée des matrices

$$E_{ij}^{(0)} = \begin{cases} E_{ij} & \text{si } i \neq j, \\ E_{ii} - E_{nn} & \text{si } i = j, \end{cases}$$

où $i, j = 1, \ldots, n$ et $(i, j) \neq (n, n)$. Une matrice $X = \sum_{i,j=1}^{n} x_{ij} E_{ij}$ de $\mathfrak{sl}(n)$ s'exprime dans cette base par la formule

$$X = \sum_{\substack{i,j=1 \\ (i,j) \neq (n,n)}}^{n} x_{ij} E_{ij}^{(0)}.$$

D'après la relation (1), on en déduit immédiatement que

$$(\operatorname{ad} X) E_{\alpha\beta}^{(0)} = \begin{cases} \sum_{i=1}^{n} (x_{i\alpha} E_{i\beta}^{(0)} - x_{\beta i} E_{\alpha i}^{(0)}) & \text{si } \alpha \neq \beta, \\ \sum_{\substack{i=1 \\ i \neq \alpha}}^{n-1} (x_{i\alpha} E_{i\alpha}^{(0)} - x_{\alpha i} E_{\alpha i}^{(0)}) - \sum_{i=1}^{n-1} (x_{in} E_{in}^{(0)} - x_{ni} E_{ni}^{(0)}) & \\ \qquad\qquad\qquad\qquad \text{si } \alpha = \beta. \end{cases}$$

En effectuant les calculs nécessaires, on obtient pour la fonctionnelle de Killing de l'algèbre de Lie $\mathfrak{sl}(n)$ la formule

$$t_{\mathfrak{sl}(n)}(X, Y) = 2n\,\mathrm{Tr}(XY) = 2nt(X, Y).$$

Donc, *pour l'algèbre de Lie* $\mathfrak{sl}(n)$, *la fonctionnelle de Killing* $t_{\mathfrak{sl}(n)}$ *et la fonctionnelle* t *ne diffèrent que d'un facteur multiplicatif.*

Il est aisé de voir maintenant qu'à la différence du cas précédent, *la fonctionnelle* $t_{\mathfrak{sl}(n)}$ *n'est pas dégénérée*, c'est-à-dire que $\mathfrak{sl}(n)^{\perp} = 0$. En effet, si $\mathrm{Tr}(XY) = 0$ pour toute matrice $Y \in \mathfrak{sl}(n)$, alors, en particulier, $x_{ij} = \mathrm{Tr}(XE_{ij}) = 0$ pour $i \neq j$ et $x_{ii} - x_{nn} = \mathrm{Tr}(X(E_{ii} - E_{nn})) = 0$ pour tout i. Donc, la matrice X est de la forme aE et, par suite, est nulle puisque $\mathrm{Tr}\, X = 0$.

Exemple 3. Une base de l'algèbre de Lie $\mathfrak{so}(n)$ des matrices antisymétriques d'ordre n est composée des matrices

$$E_{[i,j]} = \frac{E_{ij} - E_{ji}}{2}, \quad i < j,$$

et de plus, toujours en vertu de la formule (1), pour toute matrice $X = \sum_{i,j=1}^{n} x_{ij}E_{ij}$ de $\mathfrak{so}(n)$, on a la formule

$$(\mathrm{ad}\, X)\, E_{[\alpha,\beta]} = \sum_{i=1}^{n} (x_{i\alpha}E_{[i,\beta]} - x_{\beta i}E_{[\alpha,i]}).$$

Pour la fonctionnelle $t_{\mathfrak{so}(n)}$, on en déduit la formule

$$t_{\mathfrak{so}(n)}(XY) = (n-1)\,\mathrm{Tr}(XY) = (n-1)\,t(X, Y).$$

Donc, *pour l'algèbre de Lie* $\mathfrak{so}(n)$, *les fonctionnelles* $t_{\mathfrak{so}(n)}$ *et* t *diffèrent aussi d'un facteur multiplicatif.*

Comme $x_{ij} = \mathrm{Tr}(XE_{[i,j]})$ pour toute matrice $X = \sum_{i,j=1}^{n} x_{ij}E_{ij}$ de $\mathfrak{so}(n)$, il s'ensuit, en particulier, comme dans le cas précédent, que *la fonctionnelle* $t_{\mathfrak{so}(n)}$ *n'est pas dégénérée.*

La construction de la fonctionnelle de Killing peut être largement généralisée.

Définition 2. On appelle *représentation* ρ d'une algèbre de Lie $\mathfrak{g}$ dans un espace vectoriel $\mathcal{V}$ (dit *espace de représentation* de ρ) tout homomorphisme $\rho: \mathfrak{g} \to [\mathrm{End}\, \mathcal{V}]$ de l'algèbre $\mathfrak{g}$ dans l'algèbre de Lie $[\mathrm{End}\, \mathcal{V}]$ des opérateurs linéaires sur $\mathcal{V}$.

A noter que la donnée d'une représentation ρ d'une algèbre de Lie $\mathfrak{g}$ définit sur l'espace $\mathcal{V}$ de cette représentation une structure de

module (cf. leçon 5) sur l'algèbre de Lie $\mathfrak{g}$ (au moyen de la formule $xv = \rho(x)v$, où $x \in \mathfrak{g}$, $v \in \mathcal{V}$) et réciproquement, tout module sur $\mathfrak{g}$ est l'espace d'une représentation ρ telle que $\rho(x)v = xv$, $x \in \mathfrak{g}$, $v \in \mathcal{V}$. Donc, les notions de module sur l'algèbre de Lie $\mathfrak{g}$ et de représentation de l'algèbre de Lie $\mathfrak{g}$ sont identiques. La coexistence de ces deux notions relève de la routine.

Comme exemple de représentation citons l'homomorphisme ad. Cette représentation est dite *adjointe*.

Toute représentation ρ définit grâce à la formule

$$t_\rho(x, y) = t(\rho(x), \rho(y)) = \mathrm{Tr}(\rho(x)\rho(y))$$

une fonctionnelle bilinéaire symétrique invariante sur $\mathfrak{g}$ appelée *fonctionnelle de trace de la représentation* ρ.

Ainsi, la fonctionnelle de Killing n'est autre que la fonctionnelle de trace de la représentation adjointe

$$t_\mathfrak{g} = t_{\mathrm{ad}}.$$

La fonctionnelle de trace d une algèbre de Lie linéaire sera dans cette terminologie fonctionnelle de trace de la représentation identique

$$t = t_{\mathrm{id}}.$$

Proposition 1. *Dans toute algèbre de Lie $\mathfrak{g}$, les idéaux $\mathfrak{g}^2$ et $\mathfrak{r}$ sont orthogonaux relativement à la fonctionnelle de trace de toute représentation ρ. En particulier, ces idéaux sont orthogonaux relativement à la fonctionnelle de Killing.*

Démonstration. Soient $x, y \in \mathfrak{g}$ et $z \in \mathfrak{r}$. Il nous faut prouver que $t_\rho([x, y], z) = 0$. Mais par définition

$$t_\rho([x, y], z) = t(\rho[x, y], \rho z) = t([\rho x, \rho y], \rho z)$$

et, par suite, $t_\rho([x, y], z) = 0$, puisque $[\rho x, \rho y]$ appartient à l'idéal $(\rho\mathfrak{g})^2$ de l'algèbre de Lie linéaire $\rho\mathfrak{g}$ et ρz est son radical. □

Corollaire 1. *L'idéal $\mathfrak{g}^2$ d'une algèbre de Lie résoluble $\mathfrak{g}$ est orthogonal (relativement à la fonctionnelle de Killing) à l'algèbre $\mathfrak{g}$ tout entière:*

$$t_\mathfrak{g}(\mathfrak{g}^2, \mathfrak{g}) = 0. \quad \square$$

Il se trouve que cette condition nécessaire de résolubilité est suffisante aussi.

Proposition 2 (critère de Cartan de résolubilité). *Une algèbre de Lie $\mathfrak{g}$ est résoluble si et seulement si $t_\mathfrak{g}(\mathfrak{g}^2, \mathfrak{g}) = 0$.*

La démonstration de cette proposition utilise de nombreux faits d'algèbre linéaire par l'exposé desquels nous commençons.

Pour simplifier, on admettra provisoirement que le corps K est algébriquement clos (disons est le corps $\mathbb{C}$ des complexes).

Soit A un opérateur linéaire sur un espace vectoriel $\mathcal{V}$ de dimension finie. Réduisons cet opérateur à la forme normale de Jordan (cf. II, 16), c'est-à-dire trouvons dans $\mathcal{V}$ une base dans laquelle la matrice de l'opérateur A est somme directe de blocs de Jordan de la forme

$$\begin{pmatrix} \lambda & 1 & 0 \ldots 0 \\ 0 & \lambda & 1 \ldots 0 \\ \ldots & \ldots & \ldots \\ 0 \ldots & 0 & \lambda \; 1 \\ 0 \ldots & \ldots & \lambda \end{pmatrix}.$$

En remplaçant dans chacun de ces blocs les unités par des zéros, on obtient un opérateur diagonalisable A_d possédant les mêmes valeurs propres, donc le même polynôme caractéristique que l'opérateur A. L'opérateur $A_n = A - A_d$ s'obtient visiblement par substitution de zéros à toutes les racines λ. Cet opérateur est nilpotent et commute à l'opérateur A_d (donc, à l'opérateur A). Bien plus, il est immédiat de voir qu'il existe un polynôme $p(T)$ tel que $A_d = p(A)$ ($p(T)$ est un polynôme quelconque tel que $p(\lambda_i) = \lambda_i$ et $p^{(k)}(\lambda_i) = 0$ pour $k = 1, \ldots, n_i - 1$ pour chaque valeur propre λ_i de l'opérateur A, où n_i est la multiplicité de λ_i), et donc, un polynôme $q(T)$ tel que $A_n = q(A)$ (il suffit de poser $q(T) = T - p(T)$). Donc, tout opérateur commutant à l'opérateur A commutera aux opérateurs A_d et A_n.

Si maintenant $A = A' + A''$, où A' est un opérateur diagonalisable, A'' un opérateur nilpotent, et si A' et A'' commutent l'un à l'autre et donc, à l'opérateur A, alors ils commuteront aux opérateurs A_d et A_n et l'on aura la relation $A' - A_d = A_n - A''$. Mais il est aisé de voir que la différence de deux opérateurs diagonalisables (resp. nilpotents) permutables sera un opérateur diagonalisable (resp. nilpotent). Comme le seul opérateur diagonalisable à être nilpotent est l'opérateur nul, on en déduit que l'égalité $A' - A_d = A_n - A''$ n'est possible que si $A' = A_d$ et $A'' = A_n$.

Ceci prouve le lemme suivant:

Lemme 1. *Tout opérateur A défini sur un espace vectoriel $\mathcal{V}$ de dimension finie se représente d'une seule manière par la somme*

$$A = A_d + A_n \tag{2}$$

d'opérateurs A_d et A_n, tels que

1) *l'opérateur A_d est diagonalisable;*
2) *l'opérateur A_n est nilpotent;*
3) *les opérateurs A_d et A_n commutent l'un à l'autre et à l'opérateur A.*

Ceci étant, les opérateurs A_d et A_n sont des polynômes de l'opérateur A. □

L'opérateur A_n est appelé *partie nilpotente* de l'opérateur A, et l'opérateur A_d, *partie diagonalisable* de l'opérateur A. (N. Bourbaki appelle A_d *partie semi-simple de* A.) La décomposition (2) est dite *décomposition de Jordan de l'opérateur* A.

D'après les définitions générales, à tout opérateur linéaire $A: \mathcal{V} \to \mathcal{V}$ on peut associer un opérateur linéaire ad A défini sur l'espace vectoriel End $\mathcal{V}$ par la formule

$$(\operatorname{ad} A)X = AX - XA, \quad X \in \operatorname{End} \mathcal{V}.$$

Lemme 2. *On a les égalités*

$$(\operatorname{ad} A)_d = \operatorname{ad} A_d, \quad (\operatorname{ad} A)_n = \operatorname{ad} A_n.$$

D é m o n s t r a t i o n. Soit $e_1, \ldots, e_n$ une base de l'espace $\mathcal{V}$ dans laquelle la matrice de l'opérateur A_d est diagonale. Ceci signifie que

$$A_d = \lambda_1 E_{11} + \ldots + \lambda_n E_{nn},$$

où E_{ij} sont des opérateurs de base définis par la formule

$$E_{ij}e_k = \begin{cases} e_i & \text{si } k = j, \\ 0 & \text{si } k \neq j \end{cases}$$

(ces opérateurs forment une base dans l'espace vectoriel End $\mathcal{V}$ et leurs matrices dans la base $e_1, \ldots, e_n$ sont les matrices unités E_{ij}). Comme

$$E_{ij}E_{ab} = \begin{cases} E_{ib} & \text{pour } j = a, \\ 0 & \text{pour } j \neq a, \end{cases}$$

on a

$$(\operatorname{ad} A_d)E_{ij} = (\lambda_i - \lambda_j)E_{ij}.$$

Ceci signifie que la matrice de l'opérateur ad A_d est diagonale (ses éléments diagonaux étant $\lambda_i - \lambda_j$) dans la base $\{E_{ij}\}$. Donc, l'opérateur ad A_d est diagonalisable.

Par ailleurs, il est clair que pour tout $k \geqslant 0$ et tout $X \in \operatorname{End} \mathcal{V}$, l'opérateur $(\operatorname{ad} A_n)^k X$ est somme des opérateurs $\pm A_n^i X A_n^j$, où $i + j = k$. Donc, si $A_n^m = 0$, alors $(\operatorname{ad} A_n)^{2m-1} = 0$, c'est-à-dire que l'opérateur adA_n est nilpotent.

Enfin, étant donné que ad est un homomorphisme d'algèbres de Lie, on a $[\operatorname{ad} A_d, \operatorname{ad} A_n] = \operatorname{ad}[A_d, A_n] = 0$ et $\operatorname{ad} A_d + \operatorname{ad} A_n = \operatorname{ad}(A_d + A_n) = \operatorname{ad} A$, c'est-à-dire que les opérateurs ad A_d et ad A_n commutent et leur somme est égale à l'opérateur ad A.

Donc, les opérateurs ad A_d et ad A_n possèdent relativement à l'opérateur ad A les mêmes propriétés caractéristiques 1), 2) et 3) du lemme 1. Donc, ad $A_d = (\operatorname{ad} A)_d$ et ad $A_n = (\operatorname{ad} A)_n$. □

21—01642

Corollaire. *Les opérateurs* ad A_d *et* ad A_n *sont des polynômes de l'opérateur* ad A. □

Nous pouvons démontrer maintenant notre proposition maîtresse sur les algèbres de Lie linéaires.

Proposition 3 (critère de Cartan). *Si la fonctionnelle de trace d'une algèbre de Lie linéaire* $\mathfrak{g}$ *est identiquement nulle, cette algèbre est résoluble.*

Démonstration. Soit A un opérateur de $\mathfrak{g}^2$, et soient $\lambda_1, \ldots, \lambda_n$ ses valeurs propres répétées un nombre de fois égal à leur multiplicité. La proposition 3 résulte immédiatement du lemme suivant:

Lemme. *Pour toute application additive* $\beta: \mathbb{K} \to \mathbb{K}$ (*c'est-à-dire telle que* $\beta(a+b) = \beta a + \beta b$) *on a la relation*

$$\beta(\lambda_1)\lambda_1 + \ldots + \beta(\lambda_n)\lambda_n = 0.$$

En effet, le corps $\mathbb{K}$ est un espace vectoriel (de dimension infinie) sur le corps $\mathbb{Q}$ des nombres rationnels. Choisissons une base $\{u_i, i \in I\}$ du corps $\mathbb{K}$ sur $\mathbb{Q}$ (l'ensemble I est infini, mais cela n'est pas une entrave) et désignons par $\beta_i(u)$ la i-ième coordonnée d'un élément $u \in \mathbb{K}$ dans cette base. Comme $\mathbb{Q} \subset \mathbb{K}$, nous pouvons traiter $\beta_i: u \mapsto \beta_i(u)$ comme une application (visiblement additive) de $\mathbb{K}$ dans $\mathbb{K}$. Donc, d'après le lemme, on a la relation

$$\beta_i(\lambda_1)\lambda_1 + \ldots + \beta_i(\lambda_n)\lambda_n = 0.$$

En passant à la i-ième coordonnée dans cette égalité, on obtient

$$\beta_i(\lambda_1)^2 + \ldots + \beta_i(\lambda_n)^2 = 0,$$

c'est-à-dire (puisque les nombres $\beta_i(\lambda_1), \ldots, \beta_i(\lambda_n)$ sont tous rationnels) que

$$\beta_i(\lambda_1) = 0, \ldots, \beta_i(\lambda_n) = 0.$$

L'indice i étant arbitraire, ceci n'est possible que si $\lambda_1 = 0, \ldots, \lambda_n = 0$, c'est-à-dire si $A_d = 0$. Donc, l'opérateur $A = A_n$ est nilpotent.

Ceci prouve que l'idéal $\mathfrak{g}^2$ est une nilalgèbre de Lie. Donc, l'algèbre de Lie $\mathfrak{g}$ est résoluble, puisque l'idéal $\mathfrak{g}^2$ est nilpotent d'après le théorème d'Engel. □

Pour achever la démonstration de la proposition 3, il reste donc à prouver le lemme.

D é m o n s t r a t i o n d u l e m m e. Comme plus haut, on peut, sans restreindre la généralité, admettre que

$$A_d = \lambda_1 E_{11} + \ldots + \lambda_n E_{nn}.$$

Considérons l'opérateur

$$D = \beta(\lambda_1)E_{11} + \ldots + \beta(\lambda_n)E_{nn}.$$

Si $p(T)$ est un polynôme tel que $p(\lambda_i) = \beta(\lambda_i)$ pour tout i, alors $D = p(A_d)$ et, par suite, $D = p(q(A))$, où $q(T)$ est un polynôme tel que $q(A) = A_d$. Donc, l'opérateur D commute à l'opérateur A et, par suite, à A_n (et à A_d). Mais alors $(DA_n)^k = D^k A_n^k$ pour tout $k \geqslant 0$ et, donc, l'opérateur DA_n est nilpotent. Par conséquent, sa trace est nulle:

$$\operatorname{Tr} DA_n = 0$$

et donc,

$$\beta(\lambda_1)\lambda_1 + \ldots + \beta(\lambda_n)\lambda_n = \operatorname{Tr} DA_d = \operatorname{Tr} DA.$$

Il nous faut donc prouver simplement que $\operatorname{Tr} DA = 0$.

Comme $A \in \mathfrak{g}^2$, il existe dans l'algèbre $\mathfrak{g}$ des opérateurs $B_1, \ldots, B_s, C_1, \ldots, C_s$ tels que

$$A = [B_1, C_1] + \ldots + [B_s, C_s],$$

donc,

$$\operatorname{Tr} DA = \operatorname{Tr} D[B_1, C_1] + \ldots + \operatorname{Tr} D[B_s, C_s] = \\ = \operatorname{Tr}[D, B_1]C_1 + \ldots + \operatorname{Tr}[D, B_s]C_s.$$

Mais

$$(\operatorname{ad} D)E_{ij} = (\beta(\lambda_i) - \beta(\lambda_j))E_{ij}$$

et, par suite, $\operatorname{ad} D = g(\operatorname{ad} A_d)$, où $g(T)$ est un polynôme tel que $g(\lambda_i - \lambda_j) = \beta(\lambda_i) - \beta(\lambda_j)$ pour tous i et j (un tel polynôme existe, puisque pour $\lambda_i - \lambda_j = \lambda_a - \lambda_b$, on a

$$\beta(\lambda_i) - \beta(\lambda_j) = \beta(\lambda_i - \lambda_j) = \beta(\lambda_a - \lambda_b) = \beta(\lambda_a) - \beta(\lambda_b)).$$

Comme $\operatorname{ad} A_d$ est, on le sait, un polynôme de $\operatorname{ad} A$, ceci prouve qu'il existe un polynôme $f(T)$ tel que $\operatorname{ad} D = f(\operatorname{ad} A)$. Comme $A, B_i \in \mathfrak{g}$, il vient $[A, B_i] \in \mathfrak{g}$, c'est-à-dire que $(\operatorname{ad} A)B_i \in \mathfrak{g}$. Donc, $(\operatorname{ad} A)^m B_i \in \mathfrak{g}$ pour tout $m \geqslant 0$ et, par suite, $(\operatorname{ad} D)B_i = f(\operatorname{ad} A)B_i \in \mathfrak{g}$, c'est-à-dire que $[D, B_i] \in \mathfrak{g}$. Donc, par hypothèse

$$\operatorname{Tr}[D, B_i]C_i = t([D, B_i], C_i) = 0$$

pour tout $i = 1, \ldots, s$ et donc, $\operatorname{Tr} DA = 0$. □

A noter que *la proposition* 3 *est valable pour toute algèbre de Lie sur un corps* $\mathbb{K}$ *quelconque de caractéristique* 0 (et pas uniquement sur des corps algébriquement clos). En effet, en passant du corps $\mathbb{K}$ à sa clôture algébrique $\overline{\mathbb{K}}$, l'hypothèse de la proposition reste en vigueur et sa conclusion est valable sur $\mathbb{K}$ (une algèbre de Lie sur $\mathbb{K}$ est résoluble si elle l'est en tant qu'algèbre sur $\overline{\mathbb{K}}$). □

Corollaire 1. *Le radical* $\mathfrak{r}$ *d'une algèbre de Lie* $\mathfrak{g}$ *est un annulateur de l'idéal* $\mathfrak{g}^2$ *relativement à la fonctionnelle de Killing*:

$$\mathfrak{r} = (\mathfrak{g}^2)^{\perp}.$$

D é m o n s t r a t i o n. Nous savons déjà que $t_{\mathfrak{g}}(\mathfrak{g}^2, \mathfrak{r}) = 0$, c'est-à-dire que $\mathfrak{r} \subset (\mathfrak{g}^2)^{\perp}$. Par ailleurs, l'image $\mathfrak{s}$ de l'idéal $(\mathfrak{g}^2)^{\perp}$ dans l'algèbre adjointe ad $\mathfrak{g}$ est telle que la fonctionnelle t est nulle sur $\mathfrak{s}^2$. Donc, l'idéal $\mathfrak{s}^2$ est résoluble, et, par suite, il en est de même de l'idéal $(\mathfrak{g}^2)^{\perp}$. Donc, $(\mathfrak{g}^2)^{\perp} \subset \mathfrak{r}$. □

Corollaire 2. *L'algèbre de Lie* $\mathfrak{g}$ *est résoluble si*

$$t_{\mathfrak{g}}(\mathfrak{g}^2, \mathfrak{g}^2) = 0, \tag{3}$$

ou, ce qui est équivalent, si

$$\operatorname{Tr}(\operatorname{ad} x)^2 = 0 \tag{4}$$

pour tout élément $x \in \mathfrak{g}^2$.

D é m o n s t r a t i o n. D'après l'identité

$$\operatorname{Tr}(\operatorname{ad} x + \operatorname{ad} y)^2 = \operatorname{Tr}(\operatorname{ad} x)^2 + 2 \operatorname{Tr}(\operatorname{ad} x \operatorname{ad} y) + \operatorname{Tr}(\operatorname{ad} y)^2,$$

la condition (4) est équivalente à l'égalité $\operatorname{Tr}(\operatorname{ad} x \operatorname{ad} y) = 0$ pour tous éléments $x, y \in \mathfrak{g}^2$, égalité qui exprime que la fonctionnelle de trace de l'algèbre de Lie linéaire ad $\mathfrak{g}^2$ est identiquement nulle. Ce qui équivaut, par définition, à la condition (3). Donc, si ces conditions sont remplies, l'algèbre ad $\mathfrak{g}^2$ est résoluble. Il en est alors de même de l'algèbre $\mathfrak{g}$, puisque $\mathfrak{g}/\mathfrak{z} \approx \operatorname{ad} \mathfrak{g}$ et l'algèbre $\operatorname{ad} \mathfrak{g}/\operatorname{ad} \mathfrak{g}^2$ est abélienne. □

La condition $x \in \mathfrak{g}^2$ est essentielle dans ce corollaire.

Exemple 4. Considérons une algèbre de Lie $\mathfrak{g}$ à trois dimensions admettant pour base e_1, e_2, e_3, et munie d'une multiplication définie par les formules

$$[e_1, e_2] = 0, \quad [e_1, e_3] = ae_1 + be_2, \quad [e_2, e_3] = ce_1 + de_2,$$

où $ad - bc \neq 0$ et $a^2 + d^2 + 2bc \neq 0$. Pour cette algèbre de Lie, l'idéal $\mathfrak{g}^2$ est l'enveloppe linéaire des éléments e_1 et e_2, donc est une algèbre de Lie abélienne. Donc, l'algèbre de Lie $\mathfrak{g}$ est résoluble.

Par ailleurs, pour tout élément $x = x_1e_1 + x_2e_2 + x_3e_3$ de cette algèbre, l'opérateur ad x admet (dans la base e_1, e_2, e_3) la matrice

$$\begin{pmatrix} -ax_3 & -cx_3 & ax_1 + cx_2 \\ -bx_3 & -dx_3 & bx_1 + dx_2 \\ 0 & 0 & 0 \end{pmatrix}$$

et donc, l'opérateur $(\text{ad}\, x)^2$, la matrice

$$\begin{pmatrix} (a^2+bc)\,x_3^2 & (a+d)\,cx_3^2 & -(a^2+cb)\,x_1x_3-(a+d)\,cx_2x_3 \\ (a+d)\,bx_3^2 & (bc+d^2)\,x_3^2 & -(a+d)\,bx_1x_3-(bc+d^2)\,x_2x_3 \\ 0 & 0 & 0 \end{pmatrix}$$

dont la trace $(a^2 + d^2 + 2bc)x_3^2$ est nulle si seulement $x_3 = 0$ (c'est-à-dire pour $x \in \mathfrak{g}^2$).

Nous pouvons maintenant démontrer la proposition 2.

D é m o n s t r a t i o n d e l a p r o p o s i t i o n 2. Il suffit de remarquer que si $t_{\mathfrak{g}}(\mathfrak{g}^2, \mathfrak{g}) = 0$, alors *a fortiori* $t_{\mathfrak{g}}(\mathfrak{g}^2, \mathfrak{g}^2) = 0$. □

Considérons maintenant les algèbres de Lie linéaires dont la fonctionnelle de trace n'est pas dégénérée.

Proposition 4. *Une algèbre de Lie linéaire dont la fonctionnelle t n'est pas dégénérée est réductive.*

D é m o n s t r a t i o n. Comme $t(\mathfrak{g}^2, \mathfrak{r}) = 0$ et que la fonctionnelle t est invariante, alors $t(\mathfrak{g}, [\mathfrak{g}, \mathfrak{r}]) = t(\mathfrak{g}^2, \mathfrak{r}) = 0$, d'où, à cause de la non-dégénérescence, il vient que $[\mathfrak{g}, \mathfrak{r}] = 0$, c'est-à-dire que $\mathfrak{r} \subset \mathfrak{z}$. Donc, en particulier, tout idéal abélien de l'algèbre $\mathfrak{g}$ est contenu dans $\mathfrak{z}$. Par ailleurs, comme prouvé dans la leçon précédente, tout opérateur A de $\mathfrak{z} \cap \mathfrak{g}^2$ est nilpotent. Donc, quel que soit $B \in \mathfrak{g}$, l'opérateur AB sera nilpotent (puisque $AB = BA$) et, par suite, $\text{Tr}\, AB = t(A, B) = 0$. La fonctionnelle t étant non dégénérée, ceci n'est possible que si $A = 0$. Par conséquent, $\mathfrak{z} \cap \mathfrak{g}^2 = 0$ et, par suite, l'algèbre de Lie $\mathfrak{g}$ est réductive. □

La réciproque est vraie à un « isomorphisme près ». Plus exactement, on peut montrer que *toute algèbre de Lie réductive est isomorphe à une algèbre linéaire à fonctionnelle de trace non dégénérée*, mais nous glisserons sur la démonstration de cette proposition, car nous n'en aurons pas besoin.

Une autre réciproque, partielle, de la proposition 4 est contenue dans la proposition suivante, qui au contraire nous sera très utile:

Proposition 5. *La fonctionnelle de trace d'une algèbre de Lie semi-simple linéaire $\mathfrak{g}$ est non dégénérée.*

Prouvons préalablement deux lemmes.

Lemme 3. *Soient* $\mathfrak{g}$ *une algèbre de Lie, t une fonctionnelle invariante bilinéaire symétrique sur* $\mathfrak{g}$. *Alors l'annulateur* $\mathfrak{a}^{\perp}$ *de tout idéal* $\mathfrak{a}$ *de l'algèbre* $\mathfrak{g}$ *relativement à la fonctionnelle t est aussi un idéal.*

Démonstration. Si $x \in \mathfrak{a}^{\perp}$, $y \in \mathfrak{g}$ et $z \in \mathfrak{a}$, alors

$$t([x, y], z) = t(x, [y, z]) = 0$$

et donc $[x, y] \in \mathfrak{a}^{\perp}$. □

Lemme 4. *Si* $\mathfrak{g}$ *est une algèbre de Lie semi-simple linéaire,* $\mathfrak{a}^{\perp}$ *l'annulateur d'un idéal* $\mathfrak{a}$ *de* $\mathfrak{g}$ *relativement à la fonctionnelle de trace t, alors* $\mathfrak{a} \cap \mathfrak{a}^{\perp} = 0$.

Démonstration. D'après le lemme 3, l'annulateur $\mathfrak{a}^{\perp}$, et donc l'intersection $\mathfrak{a} \cap \mathfrak{a}^{\perp}$, sont des idéaux. De plus, la fonctionnelle de trace est identiquement nulle sur l'idéal $\mathfrak{a} \cap \mathfrak{a}^{\perp}$. Donc, d'après la proposition 3, l'idéal $\mathfrak{a} \cap \mathfrak{a}^{\perp}$ est résoluble et donc nul, puisque l'algèbre $\mathfrak{g}$ est semi-simple. □

Démonstration de la proposition 5. En appliquant le lemme 2 à l'idéal $\mathfrak{a} = \mathfrak{g}$, on trouve $\mathfrak{g}^{\perp} = 0$. Or, cela signifie que la fonctionnelle t n'est pas dégénérée. □

Une représentation ρ d'une algèbre de Lie $\mathfrak{g}$ est *exacte* si elle est un monomorphisme, c'est-à-dire si elle réalise un isomorphisme entre l'algèbre $\mathfrak{g}$ et l'algèbre linéaire $\rho(\mathfrak{g})$. Comme la fonctionnelle de trace de l'algèbre $\rho(\mathfrak{g})$ n'est autre que la fonctionnelle de trace de la représentation ρ, la proposition 5 revient à affirmer que *la fonctionnelle de trace de toute représentation exacte d'une algèbre de Lie semi-simple n'est pas dégénérée.*

Cette assertion s'applique, en particulier, à la représentation adjointe ad (qui est exacte, puisque le centre d'une algèbre de Lie semi-simple est nul). Comme la fonctionnelle de trace d'une représentation adjointe est très exactement la fonctionnelle de Killing, ceci prouve que *la fonctionnelle de Killing d'une algèbre de Lie semi-simple* $\mathfrak{g}$ *n'est pas dégénérée.*

Donc, l'annulateur $\mathfrak{a}^{\perp}$ d'un idéal $\mathfrak{a}$ d'une algèbre $\mathfrak{g}$ par rapport à la fonctionnelle de Killing $t_{\mathfrak{g}}$ admet une dimension supplémentaire, c'est-à-dire que

$$\dim \mathfrak{a} + \dim \mathfrak{a}^{\perp} = \dim \mathfrak{g}.$$

Par ailleurs, comme ad réalise un isomorphisme entre l'algèbre $\mathfrak{g}$ et l'algèbre linéaire ad $\mathfrak{g}$ qui à la fonctionnelle $t_{\mathfrak{g}}$ associe la fonctionnelle t de l'algèbre ad $\mathfrak{g}$, le lemme 4 nous dit que $\mathfrak{a} \cap \mathfrak{a}^{\perp} = 0$.

Donc, $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^{\perp}$.

Formulons cette assertion sous forme de la proposition suivante:

Proposition 6. *Tout idéal* $\mathfrak{a}$ *d'une algèbre de Lie* $\mathfrak{g}$ *semi-simple est élément d'une somme directe autrement dit il existe un idéal* $\mathfrak{a}^{\perp}$

tel que

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^{\perp}.$$

L'idéal supplémentaire $\mathfrak{a}^{\perp}$ *est l'annulateur de* $\mathfrak{a}$ *relativement à la fonctionnelle de Killing.* □

Corollaire. *Tout idéal et toute algèbre quotient d'une algèbre de Lie* $\mathfrak{g}$ *semi-simple sont des algèbres de Lie semi-simples.*

Démonstration. Si l'annulateur d'un sous-espace relativement à une fonctionnelle bilinéaire non dégénérée est supplémentaire direct de ce sous-espace, alors la restriction de la fonctionnelle à ce sous-espace n'est évidemment pas dégénérée non plus. Donc, la fonctionnelle de Killing n'est dégénérée sur aucun idéal $\mathfrak{a}$ de $\mathfrak{g}$. Cela signifie que l'idéal $\mathfrak{a}$ se transforme en une algèbre de Lie ad$\mathfrak{a}$ à fonctionnelle de trace non dégénérée par l'isomorphisme ad: $\mathfrak{g} \to$ ad $\mathfrak{g}$. Donc, en vertu de la proposition 4, l'algèbre ad $\mathfrak{a}$, et donc l'idéal $\mathfrak{a}$, sont des algèbres de Lie réductives, c'est-à-dire que l'idéal $\mathfrak{a}$ est la somme directe de son centre $\mathfrak{z}$ et d'un idéal semi-simple $\mathfrak{m}$. Comme tout terme de la représentation d'un idéal par une somme directe est visiblement un idéal de l'algèbre tout entière, le centre $\mathfrak{z}$ doit être nul, puisque $\mathfrak{g}$ est semi-simple. Donc, l'idéal $\mathfrak{a} = \mathfrak{m}$ est semi-simple.

L'assertion relative aux algèbres quotients se ramène à une assertion relative aux idéaux, puisqu'une algèbre quotient par un idéal $\mathfrak{a}$ est isomorphe à l'idéal supplémentaire $\mathfrak{a}^{\perp}$. □

La fonctionnelle de Killing étant non dégénérée, on en déduit également que $\mathfrak{g} = \mathfrak{g}^2$ *pour toute algèbre de Lie* $\mathfrak{g}$ *semi-simple.* En effet, d'après le corollaire 1 de la proposition 3, l'annulateur $(\mathfrak{g}^2)^{\perp}$ de l'idéal $\mathfrak{g}^2$ relativement à la fonctionnelle de Killing est nul pour une algèbre de Lie $\mathfrak{g}$ semi-simple. Donc, $\mathfrak{g} = \mathfrak{g}^2$, puisque la fonctionnelle de Killing n'est pas dégénérée. □

La condition de non-dégénérescence de la fonctionnelle de Killing est une condition non seulement nécessaire mais suffisante aussi pour que l'algèbre de Lie $\mathfrak{g}$ soit semi-simple. Pour le prouver nous remarquerons tout d'abord que *si la fonctionnelle de trace* t_ρ *d'une représentation* ρ *d'une algèbre de Lie* $\mathfrak{g}$ *n'est pas dégénérée, la représentation* ρ *est exacte.* En effet, si $\rho a = 0$, alors $t_\rho(a, x) = \mathrm{Tr}(\rho a\, \rho x) = 0$ pour tous les $x \in \mathfrak{g}$, donc, $a = 0$. □

Si, en particulier, la fonctionnelle de Killing de $\mathfrak{g}$ n'est pas dégénérée, alors la représentation adjointe de $\mathfrak{g}$ est exacte, de sorte que l'algèbre linéaire ad $\mathfrak{g}$ est isomorphe à l'algèbre $\mathfrak{g}$ (et le centre de $\mathfrak{g}$ est trivial). La fonctionnelle de trace de l'algèbre ad $\mathfrak{g}$ n'est pas dégénérée non plus, puisque l'isomorphisme ad lui associe la fonctionnelle de Killing de l'algèbre $\mathfrak{g}$. Donc, l'algèbre linéaire ad $\mathfrak{g}$

est réductive et, de plus, elle est semi-simple, puisque son centre est trivial. Donc, l'algèbre $\mathfrak{g} \approx \mathrm{ad}\, \mathfrak{g}$ est semi-simple aussi.

Nous avons ainsi prouvé la proposition suivante:

Proposition 7 (critère de Cartan de semi-simplicité). *Une algèbre de Lie $\mathfrak{g}$ est semi-simple si et seulement si sa fonctionnelle de Killing est non dégénérée.* □

D'après cette proposition et les résultats des exemples 1, 2 et 3, *les algèbres de Lie $\mathfrak{sl}(n)$ et $\mathfrak{so}(n)$ sont semi-simples, tandis que l'algèbre de Lie $\mathfrak{gl}(n)$ ne l'est pas.* □

Soit maintenant ρ une représentation non triviale d'une algèbre $\mathfrak{g}$ semi-simple (c'est-à-dire telle que $\rho a \neq 0$ au moins pour un élément $a \in \mathfrak{g}$), et soit $\mathfrak{k}$ son noyau. Alors, en vertu de la proposition 6, il existe dans l'algèbre $\mathfrak{g}$ un idéal $\mathfrak{h} = \mathfrak{k}^{\perp}$ tel que $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{h}$. La restriction de ρ à $\mathfrak{h}$ est exacte, et comme l'idéal $\mathfrak{h}$ est semi-simple (cf. corollaire de la proposition 6), la fonctionnelle t_ρ de la représentation ρ est non dégénérée sur l'idéal $\mathfrak{h}$. Donc, toute base $e_1, \ldots, e_n$ de l'idéal $\mathfrak{h}$ (traité comme un espace vectoriel sur le corps $\mathbb{K}$) admet une base $e^1, \ldots, e^n$ t_ρ-duale telle que

$$t_\rho(e_i, e^j) = \delta_i^j, \quad i, j = 1, \ldots, n.$$

Pour tout élément $x \in \mathfrak{g}$, nous posons

$$[x, e_i] = \alpha_i^j(x) e_j, \qquad [x, e^j] = \beta_i^j(x) e_j.$$

Donc $(\alpha_i^j(x))$ est la matrice de la restriction de l'opérateur linéaire $\mathrm{ad}\, x$ au sous-espace $\mathfrak{h}$ dans la base $e_1, \ldots, e_n$ et $(\beta_i^j(x))$ sa matrice dans la base $e^1, \ldots, e^n$.

Lemme 5. *Pour tout élément $x \in \mathfrak{g}$ on a les égalités*

$$\alpha_i^j(x) + \beta_i^j(x) = 0, \quad i, j = 1, \ldots, n.$$

D é m o n s t r a t i o n. Ces égalités ne sont qu'une autre transcription de la proposition relative à l'invariance de la fonctionnelle t_ρ (antisymétrie de l'opérateur linéaire $\mathrm{ad}\, x$):

$$\alpha_i^j(x) = t_\rho([x, e_i], e^j) = -t_\rho(e_i, [x, e^j]) = -\beta_i^j(x). \quad \square$$

Lemme 6. *L'opérateur linéaire*

$$C = \rho(e_i)\rho(e^i)$$

ne dépend pas du choix de la base $e_1, \ldots, e_n$.

D é m o n s t r a t i o n. La base duale d'une base quelconque $e_{i'} = c_{i'}^i e_i$ de l'espace vectoriel $\mathfrak{h}$ s'exprime par la formule $e^{j'} =$

$= c_j^{j'} e^j$. Donc,

$$\rho(e_{i'})\rho(e^{i'}) = \delta_{j'}^{i'}\rho(e_{i'})\rho(e^{j'}) = \\ = \delta_j^{i'} c_{i'}^{i} c_j^{j'} \rho(e_i)\rho(e^j) = \\ = \delta_j^i \rho(e_i)\rho(e^j) = \\ = \rho(e_i)\rho(e^i). \quad \square$$

Définition 3. L'opérateur linéaire C s'appelle *opérateur de Casimir* de la représentation ρ.

On conviendra que $C = 0$ pour la représentation triviale $\rho = 0$.

Proposition 8. *L'opérateur de Casimir commute à tout opérateur de* $\rho(\mathfrak{g})$:

$$[C, \rho(x)] = 0 \text{ pour tout } x \in \mathfrak{g}.$$

Pour $\rho \neq 0$, *la trace* Tr C *de l'opérateur de Casimir est égale à la dimension* n *de l'idéal* $\mathfrak{h}$, *de sorte que cet opérateur est non nul (et de surcroît non nilpotent).*

Démonstration. Par définition

$$\mathrm{Tr}\, C = \mathrm{Tr}(\rho(e_i)\rho(e^i)) = t_\rho(e_i, e^i) = \delta_i^i = n.$$

Comme

$$\rho(x)C = \rho(x)\rho(e_i)\rho(e^i) = \\ = [\rho(x), \rho(e_i)]\rho(e^i) + \rho(e_i)\rho(x)\rho(e^i) = \\ = \rho([x, e_i])\rho(e^i) + \rho(e_i)\rho(x)\rho(e^i) = \\ = \alpha_i^j(x)\rho(e_j)\, \rho(e^i) + \rho(e_i)\rho(x)\rho(e^i)$$

et de façon analogue

$$C\rho(x) = -\beta_j^i(x)\rho(e_j)\rho(e^i) + \rho(e_i)\rho(x)\rho(e^i),$$

il vient que $C\rho(x) - \rho(x)C = 0$ d'après le lemme 2. $\square$

La représentation ρ est dite *irréductible* s'il en est de même de l'algèbre de Lie linéaire $\rho(\mathfrak{g})$. Une telle représentation est nécessairement non triviale (si dim $\mathcal{V} > 0$).

Corollaire. *L'opérateur de Casimir d'une représentation irréductible d'une algèbre de Lie semi-simple est inversible.*

Démonstration. L'opérateur C commutant à $\rho(x)$, son noyau est invariant par les opérateurs $\rho(x)$ et, par suite, est nul à cause de l'irréductibilité. $\square$

LEÇON 19

Cohomologies d'algèbres de Lie.— Théorème de Whitehead. —Décomposition de Fitting.— Théorème généralisé de Whitehead.— Lemmes de Whitehead.— Théorème de Weyl de réductibilité complète.— Extensions des algèbres de Lie abéliennes.

Nous entamerons cette leçon par des constructions générales dont la génèse et la signification ne peuvent être expliquées que dans le cadre de l'algèbre homologique en s'appuyant sur la théorie topologique des groupes de cohomologie de groupes de Lie. Nous le ferons brièvement dans la leçon suivante, nous contentant pour l'instant d'un exposé purement formel sans aucune motivation.

Soit $\mathfrak{g}$ une algèbre de Lie, comme toujours de dimension finie et sur un corps K de caractéristique 0, et soit $\mathcal{V}$ un module sur $\mathfrak{g}$ ayant, en tant qu'espace vectoriel sur K, une dimension finie (c'est-à-dire est l'espace d'une représentation ρ de dimension finie de l'algèbre $\mathfrak{g}$).

Définition 1. On dit qu'une fonction $u = u(x_1, \ldots, x_m)$ de m variables $x_1, \ldots, x_m \in \mathfrak{g}$ à valeurs dans le module $\mathcal{V}$ est une *cochaîne à m dimensions* de l'algèbre de Lie $\mathfrak{g}$ sur $\mathcal{V}$ si:

a) elle est antisymétrique, c'est-à-dire change de signe si l'on permute deux quelconques de ses arguments;

b) est linéaire en chaque argument (à valeurs fixes des autres).

Si $m = 1$ la condition a) est dénuée de sens, de sorte que toute application linéaire $u: \mathfrak{g} \to \mathcal{V}$ sera cochaîne, et pour $m = 0$ on admet, en vertu des conventions générales relatives aux fonctions ayant zéro arguments, que u est un élément arbitraire du module $\mathcal{V}$.

Les cochaînes à m dimensions forment visiblement un espace vectoriel $C^m(\mathfrak{g}; \mathcal{V})$.

Pour toute cochaîne $u \in C^m(\mathfrak{g}; \mathcal{V})$ et tous éléments $x_1, \ldots$

$\ldots, x_{m+1} \in \mathfrak{g}$, on pose

$$(\delta u)(x_1, \ldots, x_{m+1}) = \sum_{i=1}^{m+1} (-1)^{i+1} x_i u(x_1, \ldots, \hat{x}_i, \ldots, x_{m+1}) +$$

$$+ \sum_{i=1}^{m} \sum_{j=i+1}^{m+1} (-1)^{i+j} u([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_{m+1}),$$

où les arguments « chapeautés » doivent être omis.

Il est clair que la fonction δu définie par cette formule est linéaire en chaque argument et, comme le montre un calcul automatique, est antisymétrique, c'est-à-dire est une cochaîne à $m + 1$ dimensions. L'application

$$\delta: C^m(\mathfrak{g}; \mathcal{F}) \to C^{m+1}(\mathfrak{g}; \mathcal{F})$$

est visiblement linéaire.

Pour $m = 0$

$$(\delta u)(x) = xu,$$

pour $m = 1$

$$(\delta u)(x, y) = xu(y) - yu(x) - u([x, y]),$$

pour $m = 2$

$$(\delta u)(x, y, z) = xu(y, z) - yu(x, z) + zu(x, y) -$$
$$- u([x, y], z) + u([x, z], y) - u([y, z], x).$$

La propriété fondamentale de l'application δ est que:

$$\delta \circ \delta = 0.$$

Par exemple, pour $m = 0$

$$(\delta\delta u)(x, y) = x(yu) - y(xu) - [x, y]u = 0$$

et pour $m = 1$

$$(\delta\delta u)(x, y, z) = x(yu(z) - zu(y) - u([y, z]) -$$
$$- y(xu(z) - zu(x) - u([x, z])) +$$
$$+ z(xu(y) - yu(x) - u([x, y])) -$$
$$- [x, y]u(z) + zu([x, y]) + u([[x, y], z]) +$$
$$+ [x, z]u(y) - yu([x, z]) - u([[x, z], y]) -$$
$$- [y, z]u(x) + xu([y, z]) + u([[y, z], x]) = 0.$$

Les calculs sont fastidieux dans le cas général, mais tout à fait accessibles s'ils sont conduits avec soin. Nous les laissons au lecteur.

Définition 2. On appelle *cocycle* une cochaîne u telle que $\delta u = 0$ et *cobord* une cochaîne u de la forme δv.

Les cocycles forment un sous-espace $Z^m(\mathfrak{g}; \mathcal{V})$ de l'espace vectoriel $C^m(\mathfrak{g}; \mathcal{V})$ (le noyau de l'application $\delta: C^m(\mathfrak{g}; \mathcal{V}) \to C^{m+1}(\mathfrak{g}; \mathcal{V})$), et les cobords (pour $m > 0$), le sous-espace $B^m(\mathfrak{g}; \mathcal{V})$ de l'espace vectoriel $C^m(\mathfrak{g}; \mathcal{V})$ (l'image de l'application $\delta: C^{m-1}(\mathfrak{g}; \mathcal{V}) \to C^m(\mathfrak{g}; \mathcal{V})$). La relation $\delta \circ \delta = 0$ signifie que

$$B^m(\mathfrak{g}; \mathcal{V}) \subset Z^m(\mathfrak{g}; \mathcal{V})$$

pour tout $m > 0$, de sorte qu'est défini l'espace quotient

$$H^m(\mathfrak{g}; \mathcal{V}) = Z^m(\mathfrak{g}; \mathcal{V})/B^m(\mathfrak{g}; \mathcal{V}).$$

Pour $m = 0$, on convient que $H^0(\mathfrak{g}; \mathcal{V}) = Z^0(\mathfrak{g}; \mathcal{V})$, si bien que $H^0(\mathfrak{g}; \mathcal{V})$ n'est autre que le sous-espace du module $\mathcal{V}$ composé des éléments invariants de $\mathcal{V}$, c'est-à-dire des éléments u tels que $xu = 0$ pour tout $x \in \mathfrak{g}$.

Pour $m = 1$ les cocycles sont caractérisés par la relation

$$u([x, y]) = xu(y) - yu(x), \tag{1}$$

et pour $m = 2$, par la relation

$$u([x, y], z) + u([y, z], x) + u([z, x], y) = \\ = xu(y, z) + yu(z, x) + zu(x, y). \tag{2}$$

L'égalité $H^1(\mathfrak{g}; \mathcal{V}) = 0$ signifie que si la relation (1) est réalisée, il existe un élément $v \in \mathcal{V}$ tel que

$$u(x) = xv,$$

et l'égalité $H^2(\mathfrak{g}; \mathcal{V}) = 0$, que si la relation (2) est satisfaite, il existe une application linéaire $v: \mathfrak{g} \to \mathcal{V}$ telle que

$$u(x, y) = xv(y) - yv(x) - v([x, y]).$$

Dans la suite, nous n'aurons pratiquement besoin que des deux dernières assertions.

Supposons maintenant que l'algèbre de Lie $\mathfrak{g}$ est semi-simple. Alors est défini l'opérateur de Casimir $C = \rho(e_i)\rho(e^i)$ de la représentation ρ, où $e_1, \ldots, e_n$ est une base arbitraire de l'idéal $\mathfrak{h}$ supplémentaire du noyau de la représentation ρ, et $e^1, \ldots, e^n$, sa base duale relative à la fonctionnelle t_ρ.

Proposition 1 (théorème de Whitehead). *Si l'opérateur C est inversible, alors*

$$H^m(\mathfrak{g}; \mathcal{V}) = 0 \textit{ pour tout } m \geqslant 0.$$

D é m o n s t r a t i o n. Soit $m = 0$. Si $xv = 0$ pour tout $x \in \mathfrak{g}$, alors $Cv = 0$, et, par suite, $v = 0$. Donc, $H^0(\mathfrak{g}; \mathcal{V}) = 0$.

Soit $m > 0$, et soit $u \in Z^m(\mathfrak{g}; \mathcal{F}^{\cdot})$. Alors

$$\sum_{i=1}^{m+1} (-1)^{i+1} x_i u(x_1, \ldots, \hat{x}_i, \ldots, x_{m+1}) + \sum_{i=1}^{m} \sum_{j=i+1}^{m+1} (-1)^{i+j} u([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, \hat{x}_{m+1}) = 0$$

pour tous éléments $x_1, \ldots, x_m, x_{m+1} \in \mathfrak{g}$. En remplaçant x_{m+1} par l'élément de base e_k de l'idéal $\mathfrak{h}$, en multipliant par e^k et en sommant sur k, on obtient l'identité

$$(3) \quad \sum_{i=1}^{m} (-1)^{i+1} e^k(x_i u(x_1, \ldots, \hat{x}_i, \ldots, x_m, e_k)) + \\ + (-1)^{n+2} e^k(e_k u(x_1, \ldots, x_n)) + \\ + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (-1)^{i+j} e^k u([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_m, e_k) + \\ + \sum_{i=1}^{m} (-1)^{i+m+1} e^k u([x_k, e_k], x_1, \ldots, \hat{x}_i, \ldots, x_m) = 0,$$

qui est satisfaite pour tous $x_1, \ldots, x_m \in \mathfrak{g}$.

Comme $xu = \rho(x)u$, le terme $(-1)^{n+2}e^k(e_k u(x_1, \ldots, x_m))$ de l'identité (3) n'est autre que $(-1)^{n+2}Cu(x_1, \ldots, x_m)$, et puisque $e(xu) = [e, x]u + x(eu)$, la première somme de cette identité est égale à

$$(4) \quad \sum_{i=1}^{m} (-1)^{i+1}[e^k, x_i] u(x_1, \ldots, \hat{x}_i, \ldots, x_m, e_k) + \\ + \sum_{i=1}^{m} (-1)^{i+1} x_i (e^k u(x_1, \ldots, \hat{x}_i, \ldots, x_m, e_k)) = \\ = \sum_{i=1}^{m} (-1)^i \beta_l^k(x_i) e^l u(x_1, \ldots, \hat{x}_i, \ldots, x_m, e_k) + \\ + \sum_{i=1}^{m} (-1)^{i+1} x_i v(x_1, \ldots, \hat{x}_i, \ldots, x_m),$$

où, comme dans le lemme 2 de la leçon précédente,

$$\beta_l^k(x) = t_\rho([x, e^k], e_l),$$

et

$$v(y_1, \ldots, y_{m-1}) = e^k u(y_1, \ldots, y_{m-1}, e_k)$$

pour tous $y_1, \ldots, y_{m-1} \in \mathfrak{g}$.

La dernière somme de l'identité (3) est égale à

$$\sum_{i=1} (-1)^{i+m+1} \alpha_k^l(x_i)\, e^k u(e_l, x_1, \ldots, \hat{x}_i, \ldots, x_m) =$$

$$= \sum_{i=1}^{m} (-1)^i \alpha_l^k(x_i)\, e^l u(x_1, \ldots, \hat{x}_i, \ldots, x_m, e_k),$$

où $\alpha_k^l(x) = t_\rho([x, e_k], e^l)$, et d'après le lemme 2 de la leçon précédente se simplifie avec la première somme de (4).

Comme la double somme de (3) peut être mise sous la forme

$$\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (-1)^{i+j} v([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_m),$$

ceci prouve que

$$(-1)^{n+2} Cu(x_1, \ldots, x_m) + \sum_{i=1}^{m} (-1)^{i+1} x_i v(x_1, \ldots, \hat{x}_i, \ldots, x_m) + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (-1)^{i+j} v([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_m) = 0,$$

c'est-à-dire que

$$(-1)^{n+2} Cu(x_1, \ldots, x_m) + (\delta v)(x_1, \ldots, x_m) = 0.$$

De là il s'ensuit qu'en posant

$$w(x_1, \ldots, x_m) = (-1)^{n+1} C^{-1} v(x_1, \ldots, x_m),$$

on obtient une cochaîne $w \in C^{m-1}(\mathfrak{g}; \mathcal{V})$ telle que $u = \delta w$.

Donc, chaque cocycle de dimension m de l'algèbre de Lie $\mathfrak{g}$ sur $\mathcal{V}$ est un cobord, et, par suite, $H^m(\mathfrak{g}; \mathcal{V}) = 0$. □

On rappelle qu'un opérateur linéaire C défini sur un espace vectoriel $\mathcal{V}$ est dit somme directe $A \oplus B$ d'un opérateur A défini sur un sous-espace $\mathcal{P} \subset \mathcal{V}$ et d'un opérateur B défini sur un sous-espace $\mathcal{Q} \subset \mathcal{V}$ si $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$, les sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ sont invariants par C et les opérateurs induits dans $\mathcal{P}$ et $\mathcal{Q}$ par C sont confondus respectivement avec les opérateurs A et B. Sous une forme conventionnelle, mais suggestive

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

(cf. II. 14).

Pour étudier les espaces vectoriels $H^m(\mathfrak{g}; \mathcal{V})$ dans le cas où l'opérateur de Casimir C n'est pas inversible, on se servira du lemme suivant:

Lemme 1. *Tout opérateur linéaire C défini sur un espace vectoriel $\mathcal{V}$ de dimension finie se décompose d'une seule manière en la somme directe*

$$C = A \oplus B \tag{5}$$

d'un opérateur A inversible et d'un opérateur B nilpotent.

D é m o n s t r a t i o n. L'espace $\mathcal{V}$ étant de dimension finie, la suite strictement décroissante de sous-espaces

$$\mathcal{V} \supset \operatorname{Im} C \supset \operatorname{Im} C^2 \supset \ldots \supset \operatorname{Im} C^k \supset \ldots \tag{6}$$

se stabilise, c'est-à-dire qu'il existe un k tel que $\operatorname{Im} C^k = \operatorname{Im} C^{k+1}$. Posons $\mathcal{P} = \operatorname{Im} C^k$. Il est clair que $\mathcal{P}$ est invariant par C et l'opérateur induit $A : \mathcal{P} \to \mathcal{P}$ est inversible (puisque surjectif).

La suite strictement croissante de sous-espaces

$$0 \subset \operatorname{Ker} C \subset \operatorname{Ker} C^2 \subset \ldots \subset \operatorname{Ker} C^l \subset \ldots \tag{7}$$

se stabilise pour les mêmes raisons, c'est-à-dire qu'il existe un l tel que $\operatorname{Ker} C^l = \operatorname{Ker} C^{l+1}$. Le sous-espace $\mathcal{Q} = \operatorname{Ker} C^l$ est invariant par l'opérateur C et l'opérateur induit $B: \mathcal{Q} \to \mathcal{Q}$ est nilpotent (puisque $B^l = 0$). Ceci étant, quitte à remplacer k ou l par un plus grand indice, on peut admettre que $k = l$.

Donc, pour prouver l'existence de la décomposition (5) il faut seulement établir que $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$. Par hypothèse, pour tout $v \in \mathcal{V}$, il existe un vecteur v_1 tel que le vecteur $C^k v = C^{2k} v_1$ et, par suite, le vecteur $v - C^k v_1$ appartient à $\mathcal{Q}$. Comme $C^k v_1 \in \mathcal{P}$, ceci prouve que $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$. Si $v \in \mathcal{P} \cap \mathcal{Q}$, c'est-à-dire si $v = C^k w$ et $C^k v = 0$, alors $C^{2k} w = 0$ et donc, $w \in \mathcal{Q}$. Par conséquent, $C^k w = 0$, c'est-à-dire que $v = 0$.

Ceci achève la démonstration de l'existence de la décomposition (5).

L'unicité de la décomposition (5) résulte immédiatement du fait que les sous-espaces correspondants $\mathcal{P}$ et $\mathcal{Q}$ sont les seuls sous-espaces sur lesquels se stabilisent les suites (6) et (7). □

La décomposition (5) s'appelle *décomposition de Fitting* de l'opérateur C.

Si maintenant $\mathcal{V}$ est encore l'espace de la représentation ρ d'une algèbre de Lie semi-simple $\mathfrak{g}$, et C son opérateur de Casimir, les sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ sont invariants par tous les opérateurs $\rho(x)$, $x \in \mathfrak{g}$, c'est-à-dire *sont des sous-modules*. En effet, supposons par exemple que $v \in \mathcal{P}$. Par hypothèse, il existe un vecteur $w \in \mathcal{V}$ tel que $v = C^k w$. Supposons que $\rho(x)w = w_1 + w_2$, où $w_1 \in \mathcal{P}$ et $w_2 \in \mathcal{Q}$. Comme $C^k w_2 = 0$, il vient

$$\rho(x)v = \rho(x)C^k w = C^k \rho(x) w = C^k w_1 \in \mathcal{P}.$$

De façon analogue, si $v \in \mathcal{Q}$ et $C^k v = 0$, alors $C^k\rho(x)v = \rho(x)C^k v =$ $= 0$ et, par suite, $\rho(x)v \in \mathcal{Q}$. □

Etant des modules sur l'algèbre $\mathfrak{g}$, les espaces vectoriels $\mathcal{P}$ et $\mathcal{Q}$ seront les espaces des représentations σ et τ de cette algèbre (on dit encore que la représentation ρ *est décomposée en somme directe* des représentations σ et τ). Ceci étant, il est clair que les opérateurs A et B de la décomposition de Fitting de l'opérateur de Casimir C de la représentation ρ seront les opérateurs de Casimir des représentations σ et τ. Donc, en vertu de la proposition 1, pour tout $m \geqslant 0$, on a l'égalité

$$H^m(\mathfrak{g}; \mathcal{P}) = 0.$$

D'autre part, l'opérateur de Casimir B de la représentation τ étant nilpotent, cette représentation est triviale (et $B = 0$). Cela signifie que $\mathcal{Q}$ est somme directe de sous-espaces invariants à une dimension sur lesquels l'algèbre $\mathfrak{g}$ opère trivialement. En d'autres termes, le $\mathfrak{g}$-module $\mathcal{Q}$ est somme directe de modules isomorphes au corps $\mathbb{K}$ sur lesquels l'algèbre $\mathfrak{g}$ opère trivialement. Le nombre des termes de cette somme s'appelle *multiplicité* avec laquelle $\mathbb{K}$ figure dans $\mathcal{V}$.

Mais il est évident que si le $\mathfrak{g}$-module $\mathcal{V}$ est somme directe de $\mathfrak{g}$-modules $\mathcal{V}_1, \ldots, \mathcal{V}_s$, alors pour tout $m \geqslant 0$, l'espace vectoriel $H^m(\mathfrak{g}; \mathcal{V})$ est somme directe des espaces vectoriels $H^m(\mathfrak{g}; \mathcal{V}_1), \ldots$ $\ldots, H^m(\mathfrak{g}; \mathcal{V}_s)$. Ceci et tout ce qui précède prouve la proposition suivante généralisant la proposition 1:

Proposition 2. *Pour tout module $\mathcal{V}$ sur une algèbre de Lie semi-simple $\mathfrak{g}$ et tout $m \geqslant 0$, l'espace vectoriel $H^m(\mathfrak{g}; \mathcal{V})$ est somme directe de k exemplaires de l'espace vectoriel $H^m(\mathfrak{g}; \mathbb{K})$, où k est la multiplicité avec laquelle $\mathbb{K}$ figure dans $\mathcal{V}$.* □

Donc, pour calculer les espaces vectoriels $H^m(\mathfrak{g}; \mathcal{V})$ il suffit de savoir calculer les espaces vectoriels $H^m(\mathfrak{g}; \mathbb{K})$.

A ce propos, il est utile d'observer que sur $\mathbb{K}$ la formule générale des cobords des cochaînes se simplifie considérablement et prend la forme

$$(\delta u)(x_1, \ldots, x_{m+1}) = \\ = \sum_{i=1}^{m} \sum_{j=i+1}^{m+1} (-1)^{i+j} u([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_{m+1}).$$

Par exemple, pour $m = 0$

$$(\delta u)(x) = 0,$$

pour $m = 1$

$$(\delta u)(x, y) = -u([x, y]),$$

pour $m = 2$

$$(\delta u)(x, y, z) = -u([x, y], z) + u([x, z], y) - u([y, z], x).$$

Donc $H^0(\mathfrak{g}; K) = K$ et $H^1(\mathfrak{g}; K)$ est l'espace des fonctionnelles linéaires $\mathfrak{g} \to K$ nulles sur $\mathfrak{g}^2$ (c'est-à-dire est l'annulateur Ann $\mathfrak{g}^2$ de l'espace $\mathfrak{g}^2$ sur l'espace dual $\mathfrak{g}'$).

Mais dans la leçon précédente on a prouvé qu'une algèbre de Lie semi-simple $\mathfrak{g}$ est confondue avec son idéal $\mathfrak{g}^2$. Donc, $H^1(\mathfrak{g}; K) = 0$ pour toute algèbre de Lie semi-simple $\mathfrak{g}$.

En vertu de la proposition 2, ceci prouve le

Corollaire 1 (premier lemme de Whitehead). *Pour tout module $\mathcal{V}$ sur une algèbre de Lie semi-simple $\mathfrak{g}$, on a l'égalité*

$$H^1(\mathfrak{g}; \mathcal{V}) = 0. \quad \square$$

Etablissons un résultat analogue pour $m = 2$. A cet effet, posons

$$x\xi = -\xi \circ \operatorname{ad} x$$

pour tout élément $x \in \mathfrak{g}$ et toute fonctionnelle $\xi: \mathfrak{g} \to K$. Une vérification immédiate montre que la correspondance $(x, \xi) \mapsto x\xi$ définit une structure de $\mathfrak{g}$-module sur l'espace dual $\mathfrak{g}'$. Il vaut donc la peine d'étudier les cochaînes de l'algèbre $\mathfrak{g}$ sur le module $\mathfrak{g}'$.

Pour tout $m \geqslant 1$, définissons l'application

$$\varphi: C^m(\mathfrak{g}; K) \to C^{m-1}(\mathfrak{g}; \mathfrak{g}'),$$

en posant

$$(\varphi u)(x_1, \ldots, x_{m-1})(x) = \\ = u(x_1, \ldots, x_{m-1}, x), \qquad x_1, \ldots, x_{m-1}, \quad x \in \mathfrak{g},$$

pour toute cochaîne $u \in C^m(\mathfrak{g}; K)$. Comme

$$[\delta(\varphi u)(x_1, \ldots, x_m)](x) = \\ = \sum_{i=1}^{m} (-1)^{i+1} [x_i((\varphi u)(x_1, \ldots, \hat{x}_i, \ldots, x_m))](x) + \\ + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (-1)^{i+j} [(\varphi u)([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_m)](x) = \\ = \sum_{i=1}^{m} (-1)^i [(\varphi u)(x_1, \ldots, \hat{x}_i, \ldots, x_m)]([x_i, x]) + \\ + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (-1)^{i+j} u([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_m, x) =$$

$$= \sum_{i=1}^{m} (-1)^i u(x_1, \ldots, \hat{x}_i, \ldots, x_m, [x_i, x]) +$$

$$+ \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (-1)^{i+j} u([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_m, x) =$$

$$= (\delta u)(x_1, \ldots, x_m, x) = [(\varphi(\delta u))(x_1, \ldots, x_m)] (x).$$

pour toute cochaîne $u \in C^m(\mathfrak{g}; K)$ et tous éléments $x, x_1, \ldots, x_m \in \mathfrak{g}$, il vient $\delta \circ \varphi = \varphi \circ \delta$, d'où il résulte, en particulier, que pour tout cocycle u la cochaîne φu est aussi cocycle.

Pour $m = 2$, on en déduit d'après le premier lemme de Whitehead que pour tout cocycle $u \in Z^2(\mathfrak{g}; K)$ il existe une cochaîne $\xi \in C^0(\mathfrak{g}; \mathfrak{g}') = \mathfrak{g}'$ telle que $\varphi u = \delta\xi$. Mais, pour tous éléments $x, y \in \mathfrak{g}$, on aura alors l'égalité

$$\begin{aligned} u(x, y) &= ((\varphi u)(x))(y) = ((\delta\xi)(x))(y) = \\ &= (x\xi)(y) = -(\xi \circ \operatorname{ad} x)(y) = \\ &= -\xi([x, y]) = (\delta\xi)(x, y), \end{aligned}$$

qui exprime que $u = \delta\xi$. Donc, $H^2(\mathfrak{g}; K) = 0$.

D'après la proposition 2, ceci prouve le

Corollaire 2 (deuxième lemme de Whitehead). *Pour tout module $\mathcal{V}$ sur une algèbre de Lie semi-simple $\mathfrak{g}$, on a*

$$H^2(\mathfrak{g}; \mathcal{V}) = 0. \quad \square$$

Pour $m = 3$, il existe des algèbres de Lie semi-simples telles que $H^3(\mathfrak{g}; K) \neq 0$.

Les lemmes de Whitehead, malgré leur appartenance à la modeste catégorie des lemmes, n'en sont pas moins très importants, dans la mesure où ils constituent la clef de deux théorèmes capitaux de la théorie des algèbres de Lie.

On dit qu'une représentation ρ d'une algèbre de Lie $\mathfrak{g}$ est *complètement réductible* si elle est somme directe de représentations irréductibles.

Proposition 3 (théorème de Weyl). *Toute représentation ρ d'une algèbre de Lie semi-simple $\mathfrak{g}$ sur un corps K est complètement réductible.*

Nous commencerons la démonstration de cette proposition par quelques remarques simples d'algèbre linéaire.

Soient $\mathscr{V}$ un espace vectoriel et $\mathscr{P}$ un sous-espace de $\mathscr{V}$. Considérons un sous-espace $\mathscr{Q}$ supplémentaire de $\mathscr{P}$, c'est-à-dire tel que

(8) $$\mathscr{V} = \mathscr{P} \oplus \mathscr{Q}.$$

Si $u \in \mathscr{V}$ et $u = v + w$, où $v \in \mathscr{P}$ et $w \in \mathscr{Q}$, alors en posant $Pu = v$, on obtient de toute évidence un opérateur linéaire idempotent $P: \mathscr{V} \to \mathscr{V}$ tel que $\operatorname{Im} P = \mathscr{P}$. Ces opérateurs sont dits *projecteurs sur* $\mathscr{P}$. Nous voyons donc que toute décomposition (8) définit un projecteur sur $\mathscr{P}$.

Réciproquement, soit $P: \mathscr{V} \to \mathscr{V}$ un projecteur sur $\mathscr{P}$, et soit $\mathscr{Q} = \operatorname{Ker} P$. Pour tout vecteur $u \in \mathscr{V}$, le vecteur $Pu \in \mathscr{P}$ (car $Pu = P(Pu)$) et le vecteur $u - Pu \in \mathscr{Q}$ (car $P(u - Pu) = Pu - P^2u = 0$). Comme $u = Pu + (u - Pu)$, ceci prouve que $\mathscr{V} = \mathscr{P} + \mathscr{Q}$. Mais si $u \in \mathscr{P} \cap \mathscr{Q}$, alors, *primo*, $u = Pu$ (puisque $u \in \mathscr{P} = \operatorname{Im} P$, on a $u = Pv$, où $v \in \mathscr{V}$, et, par suite, $Pu = P(Pv) = Pv = u$) et, *secundo*, $Pu = 0$ (car $u \in \mathscr{Q}$), ce qui n'est possible que pour $u = 0$. Donc, $\mathscr{V} = \mathscr{P} \oplus \mathscr{Q}$.

Ceci prouve le lemme suivant:

Lemme 2. *Les sous-espaces* $\mathscr{Q}$ *supplémentaires du sous-espace* $\mathscr{P}$ *sont en correspondance biunivoque canonique avec les projecteurs sur* $\mathscr{P}$. *Cette correspondance associe à tout projecteur P son noyau* Ker P. □

Si maintenant $\mathscr{V}$ est un module sur une algèbre de Lie $\mathfrak{g}$, *les sous-espaces* $\mathscr{P} = \operatorname{Im} P$ *et* $\mathscr{Q} = \operatorname{Ker} P$ *sont sous-modules si et seulement si l'opérateur P est un homomorphisme de modules*, c'est-à-dire permute à tous les opérateurs $\rho(x)$, $x \in \mathfrak{g}$, où ρ est une représentation définie par le module $\mathscr{V}$. En effet, si P est un homomorphisme et $u \in \mathscr{P}$ (et, par suite, $u = Pu$), alors $xu = \rho(x)u = \rho(x)Pu = P\rho(x)\,u \in \mathscr{P}$ pour tout $x \in \mathfrak{g}$. De façon analogue, si $u \in \mathscr{Q}$, alors $P(xu) = \rho(x)P(u) = 0$ et donc, $xu \in \mathscr{Q}$. Réciproquement, si $\mathscr{P}$ et $\mathscr{Q}$ sont des sous-modules, alors pour tous éléments $u \in \mathscr{V}$ et $x \in \mathfrak{g}$ on a l'égalité $xu = x(Pu + (u - Pu)) = xPu + x(u - Pu)$, où $xPu \in \mathscr{P}$ et $x(u - Pu) \in \mathscr{Q}$, qui montre que $x(Pu) = P(xu)$. □

Démonstration de la proposition 3. Une récurrence évidente (utilisant le fait que la dimension de la représentation ρ est finie) montre que pour prouver la proposition 3 il suffit d'établir que pour tout sous-module $\mathscr{P}$ d'un $\mathfrak{g}$-module $\mathscr{V}$, on a la décomposition (6) dans laquelle $\mathscr{Q}$ est un sous-module, c'est-à-dire de prouver qu'*il existe un projecteur* $P: \mathscr{V} \to \mathscr{V}$ *sur* $\mathscr{P}$ *qui est un homomorphisme de modules*. Considérons à cet effet l'ensemble $\mathscr{W}$ des opérateurs linéaires $A: \mathscr{V} \to \mathscr{V}$ pour lesquels $\operatorname{Im} A \subset \mathscr{P} \subset \operatorname{Ker} A$ (et, par suite, $A^2 = 0$). Un calcul automatique montre que $\mathscr{W}$ est un sous-espace de End $\mathscr{V}$ et, de surcroît, un module sur $\mathfrak{g}$ pour l'opération

$$xA = [\rho(x),\ A], \qquad x = \mathfrak{g}, \qquad A \in \mathscr{W}.$$

Si maintenant P_0 est un projecteur sur $\mathscr{P}$, alors il est immédiat que $[\rho(x), P_0] \in \mathscr{W}$ pour tout $x \in \mathfrak{g}$, et, par suite, la formule

$$u(x) = [\rho(x), P_0], \qquad x \in \mathfrak{g},$$

définit une application linéaire $u: \mathfrak{g} \to \mathscr{W}$, c'est-à-dire une cochaîne de $C^1(\mathfrak{g}; \mathscr{W})$. Cette cochaîne est un cocycle, puisque d'après l'identité de Jacobi

$$xu(y) - yu(x) - u([x, y]) = [\rho(x), [\rho(y), P_0]] - \\ - [\rho(y), [\rho(x), P_0]] - [\rho[x, y], P_0] = 0.$$

Donc, en vertu du premier lemme de Whitehead, il existe un opérateur $A \in \mathscr{W}$ tel que $u(x) = xA$ pour tout $x \in \mathfrak{g}$, c'est-à-dire tel que $[\rho(x), P_0] = [\rho(x), A]$. Cette relation signifie que l'opérateur $P = P_0 - A$ commute à tous les opérateurs $\rho(x)$, $x \in \mathfrak{g}$, c'est-à-dire est un homomorphisme de $\mathscr{V}$ dans $\mathscr{V}$. Ceci étant, $Pv \in \mathscr{V}$ pour tout $v \in \mathscr{V}$ et $Pv = P_0v = v$ si $v \in \mathscr{P}$, de sorte que P est un projecteur sur $\mathscr{P}$. □

L'utilisation du deuxième lemme de Whitehead à des fins identiques implique quelques préparatifs.

On dira qu'une algèbre de Lie $\mathfrak{h}$ est une *extension* d'une algèbre de Lie $\mathfrak{a}$ par une algèbre de Lie $\mathfrak{g}$ si $\mathfrak{a}$ est un idéal de $\mathfrak{h}$ et si l'algèbre quotient $\mathfrak{h}/\mathfrak{a}$ est isomorphe à l'algèbre $\mathfrak{g}$. La dernière condition signifie qu'il existe un épimorphisme d'algèbres de Lie $\alpha: \mathfrak{h} \to \mathfrak{g}$ dont le noyau est l'idéal $\mathfrak{a}$. Nous admettrons que cet épimorphisme est choisi une fois pour toutes.

On dit que l'extension $\mathfrak{h}$ est *triviale* si $\mathfrak{h}$ contient une sous-algèbre qui est transformée isomorphiquement en l'algèbre $\mathfrak{g}$ par l'épimorphisme α.

Il est clair qu'on peut toujours trouver une application linéaire

$$\beta: \mathfrak{g} \to \mathfrak{h},$$

telle que $\alpha \circ \beta = \mathrm{id}_{\mathfrak{g}}$ (section de l'homomorphisme α). Alors, l'algèbre de Lie $\mathfrak{h}$ se décomposera en tant qu'espace vectoriel en somme directe de l'idéal $\mathfrak{a}$ et du sous-espace $\operatorname{Im} \beta$. L'extension $\mathfrak{h}$ est triviale si et seulement si l'application β peut être choisie parmi les homomorphismes d'algèbres de Lie.

L'écart de l'application β par rapport à un homomorphisme d'algèbres est mesuré par la fonction $u = u(x, y)$ définie par la formule

$$u(x, y) = [\beta x, \beta y] - \beta[x, y], \qquad x, y \in \mathfrak{g}.$$

Comme $\alpha u(x, y) = [\alpha\beta x, \alpha\beta y] - \alpha\beta[x, y] = 0$, la fonction u peut être considérée comme prenant ses valeurs dans l'idéal $\mathfrak{a}$. Il est évident que cette fonction est linéaire en x et en y et est antisymé-

trique, c'est-à-dire est une cochaîne à deux dimensions de l'algèbre de Lie $\mathfrak{g}$ à valeurs dans l'idéal $\mathfrak{a}$.

Du reste, nous avons brûlé une étape en faisant cette conclusion, car l'idéal $\mathfrak{a}$ n'est pas généralement un $\mathfrak{g}$-module, propriété qui est exigée de l'ensemble de valeurs des cochaînes. Pour tourner cette difficulté, on suppose que $\mathfrak{a}^2 = 0$, c'est-à-dire que *l'algèbre* $\mathfrak{a}$ *est abélienne*. Alors la formule

$$xa = [\beta x, a], \quad x \in \mathfrak{g}, \ a \in \mathfrak{a},$$

définit (indépendamment du choix de la section β) une structure unique de $\mathfrak{g}$-module sur l'idéal $\mathfrak{a}$, donc on peut appeler la fonction u cochaîne *jure et facto*.

Il est immédiat de voir maintenant qu'*une cochaîne* $u \in C^2(\mathfrak{g}; \mathfrak{a})$ *est un cocycle*. En effet, en appliquant deux fois l'identité de Jacobi, on obtient

$$\begin{aligned}
\delta u)(x, y, z) &= xu(y, z) - yu(x, z) + zu(x, y) - \\
&- u([x, y], z) + u([x, z], y) - u([y, z], x) = \\
&= [\beta x, [\beta y, \beta z] - \beta[y, z]] - [\beta y, [\beta x, \beta z] - \\
&- \beta[x, z]] + [\beta z, [\beta x, \beta y] - \beta[x, y]] - \\
&- [\beta[x, y], \beta z] + \beta[[x, y], z] + [\beta [x, z], \beta y] - \\
&- \beta[[x, z], y] - [\beta[y, z], \beta x] + \beta [[y, z], x] = \\
&= [\beta x, [\beta y, \beta z]] - [\beta y, [\beta x, \beta z]] + [\beta z, [\beta x, \beta y]] + \\
&+ \beta([[x, y], z] - [[x, z], y] + [[y, z], x]) = 0
\end{aligned}$$

quels que soient $x, y, z \in \mathfrak{g}$. □

Soit β' une autre section de l'épimorphisme α. La différence $v = \beta - \beta'$ peut être traitée comme une application linéaire $\mathfrak{g} \to \mathfrak{a}$, c'est-à-dire comme une cochaîne de $C^1(\mathfrak{g}; \mathfrak{a})$. Comme

$$\begin{aligned}
(\delta v)(x, y) &= xv(y) - yv(x) - v([x, y]) = \\
&= [\beta x, \beta y - \beta' y] - [\beta' y, \beta x - \beta' x] - \beta[x, y] + \beta'[x, y] = \\
&= ([\beta x, \beta y] - \beta[x, y]) - ([\beta' x, \beta' y] - \beta'[x, y]),
\end{aligned}$$

la substitution de β' à β implique la substitution de $u - \delta v$ à u. Donc, si $u = \delta v$, l'application $\beta' = \beta - v$ sera un homomorphisme d'algèbres de Lie, c'est-à-dire que l'extension $\mathfrak{h}$ sera triviale.

Mais, d'après le deuxième lemme de Whitehead, la condition $u = \delta v$ est automatiquement remplie si l'algèbre de Lie est semi-simple. Ceci prouve la proposition suivante:

Proposition 4. *Toute extension* $\mathfrak{h}$ *d'une algèbre de Lie abélienne* $\mathfrak{a}$ *par une algèbre de Lie semi-simple* $\mathfrak{g}$ *est triviale.* □

Corollaire. *Une algèbre de Lie $\mathfrak{g}$ est réductive si et seulement si son centre $\mathfrak{z}$ est confondu avec son radical $\mathfrak{r}$:*

$$\mathfrak{z} = \mathfrak{r}.$$

D é m o n s t r a t i o n. Dans la leçon 12, nous avons prouvé que $\mathfrak{r} = \mathfrak{z}$ pour une algèbre réductive. Supposons réciproquement que $\mathfrak{r} = \mathfrak{z}$. Alors $\mathfrak{g}$ est l'extension de l'algèbre abélienne $\mathfrak{z}$ par l'algèbre semi-simple $\mathfrak{g}/\mathfrak{r}$. Donc, en vertu de la proposition 4, on a $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{m}$, où $\mathfrak{m}$ est une sous-algèbre semi-simple. Mais comme $[\mathfrak{z}, \mathfrak{m}] = 0 \subset \mathfrak{m}$, la sous-algèbre $\mathfrak{m}$ est un idéal. Donc, l'algèbre de Lie $\mathfrak{g}$ est réductive. □

LEÇON 20

Théorème de Lévi.— Algèbres et groupes de Lie simples.— Groupes caïniens et groupes unimodulaires.— Lemme de Schur. — Centre d'un groupe de Lie matriciel simple.— Exemple de groupe de Lie non matriciel.— Cohomologies de de Rham.— Groupes de cohomologie d'une algèbre de Lie de champs de vecteurs.— Comparaison des groupes de cohomologie d'un groupe de Lie et de son algèbre de Lie.

Toute algèbre de Lie $\mathfrak{g}$ est l'extension de son radical $\mathfrak{r}$ par l'algèbre $\mathfrak{g}/\mathfrak{r}$. On dira qu'une algèbre de Lie $\mathfrak{g}$ *admet la décomposition de Lévi* si son extension est triviale, c'est-à-dire s'il existe une sous-algèbre $\mathfrak{m}$ (nécessairement semi-simple) de $\mathfrak{g}$ telle que

$$\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{m}.$$

Si l'algèbre $\mathfrak{g}$ est semi-simple ou résoluble, elle admet manifestement la décomposition de Lévi. D'après la proposition 4 de la leçon précédente, une algèbre $\mathfrak{g}$ dont le radical $\mathfrak{r}$ est abélien admet la décomposition de Lévi.

Proposition 1 (théorème de Lévi). *Toute algèbre de Lie $\mathfrak{g}$ sur un corps $\mathbb{K}$ admet la décomposition de Lévi.*

Démonstration. Procédons par récurrence sur la dimension $n = \dim \mathfrak{r}$ du radical de l'algèbre $\mathfrak{g}$. Si $n = 0$ ou $n = 1$ et si l'idéal $\mathfrak{a} = \mathfrak{r}^2$ est nul, l'algèbre $\mathfrak{g}$ admet la décomposition de Lévi d'après ce qui précède. Soit $n > 1$, et soit $\mathfrak{a} \neq 0$. Alors $\dim \mathfrak{r}/\mathfrak{a} < n$, et comme $\mathfrak{r}/\mathfrak{a}$ est le radical de l'algèbre quotient $\mathfrak{g}/\mathfrak{a}$, l'algèbre $\mathfrak{g}/\mathfrak{a}$ admet la décomposition de Lévi en vertu de l'hypothèse de récurrence. Cela signifie qu'il existe une sous-algèbre $\mathfrak{b}$ de $\mathfrak{g}$ telle que $\mathfrak{g} = \mathfrak{r} + \mathfrak{b}$ et $\mathfrak{r} \cap \mathfrak{b} = \mathfrak{a}$. L'idéal $\mathfrak{a}$ est un idéal résoluble de l'algèbre $\mathfrak{b}$ tel que l'algèbre quotient $\mathfrak{b}/\mathfrak{a}$ est semi-simple (car isomorphe à l'algèbre semi-simple $\mathfrak{g}/\mathfrak{r}$). Ceci exprime que $\mathfrak{a}$ est le radical de l'algèbre $\mathfrak{b}$, et comme $\dim \mathfrak{a} < n$, il existe d'après l'hypothèse de récur-

rence une sous-algèbre semi-simple $\mathfrak{m}$ de $\mathfrak{b}$ telle que $\mathfrak{b} = \mathfrak{a} \oplus \mathfrak{m}$. Mais alors $\mathfrak{r} \cap \mathfrak{m} = \mathfrak{r} \cap \mathfrak{b} \cap \mathfrak{m} = \mathfrak{a} \cap \mathfrak{m} = 0$ et $\mathfrak{r} + \mathfrak{m} = \mathfrak{r} + \mathfrak{r} \cap \mathfrak{b} + \mathfrak{m} = \mathfrak{r} + \mathfrak{a} + \mathfrak{m} = \mathfrak{r} + \mathfrak{b} = \mathfrak{g}$, de sorte que $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{m}$. □

Nous aurions pu nous attaquer immédiatement à la démonstration du théorème d'Ado, mais nous allons nous en éloigner et étudier sa validité pour les groupes de Lie. Nous nous pencherons également sur la corrélation de la théorie algébrique formelle des cohomologies de la leçon 19 et des théories des cohomologies connues en topologie.

Suivant la terminologie en usage en algèbre, on dira qu'une algèbre de Lie $\mathfrak{g}$ est *simple* si elle ne possède pas d'idéaux non nuls différents de $\mathfrak{g}$.

Il est clair qu'une algèbre de Lie abélienne simple non triviale est nécessairement à une dimension, et une algèbre de Lie simple non abélienne, semi-simple.

On démontre que *l'algèbre de Lie* $\mathfrak{sl}(n)$ *est simple*. Les calculs sont assez fastidieux, aussi nous bornerons-nous au seul cas $n = 2$ qui nous intéresse.

L'algèbre de Lie $\mathfrak{sl}(2)$ est d'ordre trois et les matrices

$$H = E_{11} - E_{22}, \qquad E_1 = E_{12}, \qquad E_{-1} = E_{21}$$

en sont une base. Un calcul direct montre que

$$[H, E_1] = 2E_1, \qquad [H, E_{-1}] = -2E_{-1}, \qquad [E_1, E_{-1}] = H,$$

d'où, en particulier, il résulte que l'idéal $\mathfrak{a}$ de l'algèbre $\mathfrak{sl}(2)$ qui contient l'un au moins des éléments H, E_1, E_{-1}, les contiendra tous et, par suite, sera confondu avec l'algèbre $\mathfrak{sl}$ (2) tout entière. Par ailleurs, si $A = aH + a_1E_1 + a_{-1}E_{-1} \in \mathfrak{a}$, alors $[H, A] = 2a_1E_1 - 2a_{-1}E_{-1} \in \mathfrak{a}$, d'où

$$aH + 2a_1E_1 = A + \frac{[H, A]}{2} \in \mathfrak{a},$$

et donc, $[H, aH + 2a_1E_1] = 4a_1E_1 \in \mathfrak{a}$. Par conséquent, si $\mathfrak{a} \neq \mathfrak{sl}(n)$, alors $a_1 = 0$ et, par suite, $aH \in \mathfrak{a}$, ce qui n'est possible que pour $a = 0$. Donc, $A = a_{-1}E_{-1} \in \mathfrak{a}$, et l'on déduit encore que ou bien $\mathfrak{a} = \mathfrak{sl}(n)$, ou bien $a_{-1} = 0$. En définitive, si $\mathfrak{a} \neq \mathfrak{sl}(n)$, alors $A = 0$, c'est-à-dire que $\mathfrak{a} = 0$. □

Un groupe de Lie connexe G est *simple* si son algèbre de Lie est non abélienne et simple, ou, ce qui est équivalent, si tout sous-groupe invariant de G, différent de G, est de dimension nulle (et, par suite, si fermé, alors discret). Donc, si un groupe G est simple, tout épimorphisme $G \to G'$ soit est un revêtement de groupe (un isomorphisme local), soit le groupe G' est trivial (est composé uniquement de l'unité). Sous ces conditions, le groupe G' est simple aussi, puis-

que par définition, un groupe de Lie est simple ou non en même temps que les groupes qui lui sont localement isomorphes.

Un groupe de Lie simple est le groupe SL(2) ainsi que son revêtement universel $\widetilde{SL(2)}$.

De façon analogue, un groupe de Lie connexe est *semi-simple* si son algèbre de Lie l'est.

A noter que tout groupe de Lie simple est semi-simple.

On rappelle que le *groupe des commutateurs* $[G, G]$ d'un groupe G est le sous-groupe de G engendré par les éléments de la forme $aba^{-1}b^{-1}$. Un groupe G sera dit *caïnien* *) si $[G, G] = G$. De la proposition 2 de la leçon 4, il résulte immédiatement qu'*un groupe de Lie connexe G est groupe caïnien si et seulement si son algèbre de Lie $\mathfrak{g}$ est telle que* $\mathfrak{g} = \mathfrak{g}^2$.

On en déduit, en particulier, que *tout groupe de Lie semi-simple est caïnien.*

Sont des groupes de Lie caïniens le groupe de Lie SL(2) ainsi que tout groupe de Lie localement isomorphe à SL(2), par exemple le groupe $\widetilde{SL(2)}$.

Le théorème du déterminant du produit de matrices nous dit que *tout groupe de Lie matriciel caïnien* (*et, en particulier, semi-simple*) *est unimodulaire*, c'est-à-dire est composé de matrices unimodulaires.

Il est évident par ailleurs que *tout groupe quotient d'un groupe caïnien est caïnien.*

Soit $\mathfrak{g}$ une algèbre de Lie linéaire sur C, composée des opérateurs linéaires de $\mathcal{V}$ dans $\mathcal{V}$, où $\mathcal{V}$ est un espace vectoriel complexe de dimension finie. Si un opérateur linéaire A (n'appartenant pas en général à $\mathfrak{g}$) commute à tous les opérateurs de $\mathfrak{g}$, alors Ker A et Im A sont invariants par ces opérateurs, et, par suite, si $\mathfrak{g}$ est irréductible, alors ou bien $A = 0$, ou bien A est non dégénéré. Or, pour toute valeur propre λ de l'opérateur A, l'opérateur $A - \lambda E$ est dégénéré et commute aussi à tous les opérateurs de $\mathfrak{g}$. Donc, $A - \lambda E = 0$. Ceci prouve que *les seuls opérateurs commutant à tous les éléments d'une algèbre de Lie linéaire irréductible $\mathfrak{g}$ sur C sont les opérateurs scalaires* λE.

Cette assertion s'appelle *lemme de Schur*. A noter que pour la prouver nous n'avons à aucun moment sollicité la structure algébrique de l'algèbre de Lie $\mathfrak{g}$. Donc, le lemme de Schur est valable pour toutes familles $\mathfrak{g}$ irréductibles (c'est-à-dire ne possédant pas de sous-espaces invariants non triviaux) d'opérateurs linéaires sur C.

---

*) Ce groupe doit son nom au fait que la structure de groupe abélien y est annihilée. Comme quoi l'histoire se répète même en mathématiques abstraites.

Considérons maintenant une algèbre de Lie linéaire $\mathfrak{g}$ sur $\mathbb{R}$, composée d'opérateurs linéaires de $\mathcal{V}$ dans $\mathcal{V}$, où $\mathcal{V}$ est un espace vectoriel réel. Pour utiliser le lemme de Schur, passons de l'algèbre $\mathfrak{g}$ à sa complexifiée $\mathfrak{g}^{\mathbb{C}}$ qui est composée des matrices (des opérateurs linéaires) $A + iB$, A, $B \in \mathfrak{g}$, et opère dans le complexifié $\mathcal{V}^{\mathbb{C}}$ de $\mathcal{V}$. Il est évident que si l'algèbre de Lie $\mathfrak{g}$ (sur $\mathbb{R}$) est semi-simple, l'algèbre de Lie $\mathfrak{g}^{\mathbb{C}}$ (sur $\mathbb{C}$) le sera aussi. Mais si $\mathfrak{g}^{\mathbb{C}}$ est semi-simple, le théorème de Weyl (proposition 3 de la leçon 19) nous dit que l'espace $\mathcal{V}^{\mathbb{C}}$ se décompose en somme directe de sous-espaces invariants $\mathcal{V}_i$, $i = 1, \ldots, s$, sur chacun desquels l'algèbre $\mathfrak{g}^{\mathbb{C}}$ opère irréductiblement. Dans le langage matriciel, cela signifie que chaque matrice de $\mathfrak{g}^{\mathbb{C}}$ (donc, de $\mathfrak{g}$) est, dans une base dûment choisie, de la forme

$$A_1 \oplus \ldots \oplus A_s, \tag{1}$$

et, de plus, pour tout $i = 1, \ldots, s$, l'application $\rho_i\colon A \mapsto A_i$ sera une représentation irréductible de l'algèbre $\mathfrak{g}^{\mathbb{C}}$ dans l'espace $\mathcal{V}_i$.

Sans perdre ceci de vue, supposons que l'algèbre $\mathfrak{g}$ est l'algèbre de Lie d'un groupe de Lie G linéaire (= matriciel) semi-simple (donc, connexe) opérant sur l'espace $\mathcal{V}$. Le groupe G aussi opère canoniquement sur l'espace $\mathcal{V}^{\mathbb{C}}$, et, de plus, les sous-espaces $\mathcal{V}_i$ seront tous invariants par cette action (ils sont invariants par les opérateurs $A \in \mathfrak{g} = \operatorname{Re} \mathfrak{g}^{\mathbb{C}}$, donc par les opérateurs e^{tA}, $t \in \mathbb{R}$, qui engendrent — en vertu de la connexité — le groupe G). Donc, tout élément du groupe G admet aussi la décomposition (1) (en général avec des matrices $A_1, \ldots, A_s$ complexes). Dans ces conditions, les matrices A_i formeront un groupe G_i (généralement réductible et éventuellement trivial) opérant sur l'espace $\mathcal{V}_i$ et image épimorphique du groupe G.

Signalons que *tous les groupes G_i sont unimodulaires*. En effet, étant semi-simple, le groupe G est caïnien. Donc, tous les groupes G_i sont caïniens et, par suite, unimodulaires. □

Soit maintenant A un élément du centre du groupe G. Etant permutable à tout élément de G, l'opérateur A l'est à tout élément de l'algèbre de Lie $\mathfrak{g}$, et, par suite, à tout élément de l'algèbre de Lie $\mathfrak{g}^{\mathbb{C}}$. Donc, tout opérateur $A_i\colon \mathcal{V}_i \to \mathcal{V}_i$, $i = 1, \ldots, s$, de la décomposition (1) de l'opérateur A commute à chaque élément de l'algèbre irréductible $\rho_i(\mathfrak{g}^{\mathbb{C}})$ et, par suite, est, en vertu de lemme de Schur, un opérateur scalaire de la forme $\lambda_i E$, où $\lambda_i \in \mathbb{C}$. Par ailleurs, l'opérateur A_i est unimodulaire en tant qu'élément du groupe G_i. Donc, $\lambda_i^{n_i} = 1$, où $n_i = \dim \mathcal{V}_i$ et, par suite, l'opérateur A appartient à un ensemble fini (contenant $\leqslant n_1 n_2 \ldots n_s$ éléments). Ceci prouve la proposition suivante qui était le principal objectif de nos raisonnements:

Proposition 1. *Le centre de tout groupe de Lie matriciel semi-simple (et, en particulier, simple) est fini.* □

Cette proposition nous permet de prouver sans difficultés que le théorème d'Ado n'est pas valable pour les groupes de Lie, c'est-à-dire qu'*il existe des groupes de Lie connexes qui ne sont isomorphes à aucun groupe matriciel.* Il suffit à cet effet de trouver un groupe de Lie simple connexe ayant un centre infini. Nous allons montrer qu'un tel groupe est le groupe de revêtement universel $\widetilde{SL(2)}$ du groupe $SL(2)$ des matrices unimodulaires d'ordre deux (dont on sait déjà qu'il est simple). Pour cela il nous faut expliciter le groupe $\widetilde{SL(2)}$.

En anticipant un peu, désignons par $\widetilde{SL(2)}$ la variété différentiable, produit $\mathbb{R} \times D$ de la droite réelle $\mathbb{R}$ par le disque unité D du plan des nombres complexes z tels que $|z| < 1$.

Il est immédiat de voir que l'application définie par la formule $z \mapsto \frac{1+z}{1+\bar{z}}$ envoie le disque D dans le cercle $|w| = 1$ privé du point $w = -1$. Donc, pour tout point $z \in D$, il existe un seul nombre $t \in]-\pi, \pi[$ tel que

$$e^{it} = \frac{1+z}{1+\bar{z}}.$$

Désignons ce nombre t par $\frac{1}{i} \ln \frac{1+z}{1+\bar{z}}$.

Munissons l'ensemble $\widetilde{SL(2)}$ de la multiplication

$$(1) \qquad (x, u)(y, v) = \left(x + y + t, \ \frac{u + e^{2iy}v}{e^{2iy} + u\bar{v}}\right), \quad x, y \in \mathbb{R}, \quad u, v \in D,$$

où

$$t = \frac{1}{2i} \ln \frac{1 + e^{-2iy}u\bar{v}}{1 + e^{2iy}\bar{u}v}.$$

Comme $|e^{2iy} + u\bar{v}|^2 - |u + e^{2iy}v|^2 = (1 - |u|^2)(1 - |v|^2) > 0$, il vient

$$\left|\frac{u + e^{2iy}v}{e^{2iy} + u\bar{v}}\right| < 1,$$

donc la formule (1) définit de façon unique la multiplication sur $\widetilde{SL(2)}$.

Cette multiplication admet $(0, 0)$ pour élément neutre, et pour tout élément (x, u) existe l'élément inverse

$$(x, u)^{-1} = (-x, -e^{2ix}u).$$

Par ailleurs, un calcul direct, mais assez long, montre que la multiplication (1) est associative. Cette multiplication étant visiblement différentiable, ceci prouve qu'*elle munit la variété* $\widetilde{\mathrm{SL}(2)}$ *d'une structure de groupe de Lie.*

Le centre de ce groupe est formé des éléments (x, u) tels que

$$\frac{u+e^{2iy}v}{e^{2iy}+u\bar{v}}=\frac{v+e^{2ix}u}{e^{2ix}+v\bar{u}}$$

identiquement en y et en v. En particulier, pour $v = 0$ l'égalité $u = e^{2iy}u$ doit avoir lieu identiquement en y, ce qui n'est possible que si $u = 0$. Donc, $v = e^{2ix}v$ identiquement en v, ce qui n'est possible que pour $x = \pi n$, où n est un entier arbitraire. Réciproquement, il est clair que tous les éléments de la forme $(\pi n, 0)$ appartiennent au centre du groupe $\widetilde{\mathrm{SL}(2)}$. Donc, *le centre du groupe* $\widetilde{\mathrm{SL}(2)}$ *est infini.*

Reste maintenant à établir que le groupe $\widetilde{\mathrm{SL}(2)}$ est bien un revêtement universel du groupe $\mathrm{SL}(2)$. Considérons à cet effet l'application $\widetilde{\mathrm{SL}(2)} \to \mathrm{SL}(2)$ qui à tout élément $(x, u) \in \widetilde{\mathrm{SL}(2)}$ associe la matrice

$$\frac{1}{\sqrt{1-|u|^2}}\begin{pmatrix}\cos x+|u|\cos(x+\varphi) & |u|\sin(x+\varphi)-\sin x\\ |u|\sin(x+\varphi)+\sin x & \cos x-|u|\cos(x+\varphi)\end{pmatrix},$$

où $\varphi = \arg u$.

Un calcul élémentaire, mais hélas assez long, montre que cette application est un homomorphisme et que son noyau est composé des éléments de la forme $(2\pi n, 0)$, donc il est discret. Comme $\dim \widetilde{\mathrm{SL}(2)} = \dim \mathrm{SL}(2) = 3$, on en déduit en vertu du lemme 1 de la leçon 13 que *l'homomorphisme* $\widetilde{\mathrm{SL}(2)} \to \mathrm{SL}(2)$ *est un revêtement de groupe.*

Le groupe $\widetilde{\mathrm{SL}(2)}$ étant de toute évidence simplement connexe (puisque la variété $\mathbb{R} \times D = \widetilde{\mathrm{SL}(2)}$ est le produit de variétés $\mathbb{R}$ et D simplement connexes), on obtient ce qu'on voulait. □

Discussion. Expliquons en conclusion l'origine de cette construction du groupe $\widetilde{\mathrm{SL}(2)}$ (bien que formellement tout ait déjà été prouvé).

D'après le théorème de décomposition polaire (cf. II, 21, proposition 4), toute matrice unimodulaire $A \in \mathrm{SL}(2)$ se décompose d'une seule façon en un produit UP d'une matrice de rotation U et d'une matrice P unimodulaire définie positive. La matrice U

est définie par l'angle de rotation x, et la matrice P, qui est de la forme

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

où $a > 0$ et $ac - b^2 = 1$, est définie par les nombres $a > 0$ et b, c'est-à-dire par le nombre complexe $z = a + ib$ qui est situé dans le demi-plan de droite $a > 0$. En passant par la transformation homographique du demi-plan de droite au disque unité, on obtient donc que chaque matrice $A \in \mathrm{SL}(2)$ est caractérisée de façon unique par le couple (x, u) (de sorte que le groupe SL(2) est difféomorphe au produit $S^1 \times D$), et, de plus, il s'avère qu'à la multiplication des matrices est associée précisément la multiplication des couples définie par la formule (1). Pour obtenir le groupe $\widetilde{\mathrm{SL}(2)}$, il reste maintenant à traiter x non pas comme un angle, mais comme un point de $\mathbb{R}$.

Remarque 1. On peut démontrer que le centre du groupe de revêtement simplement connexe $\widetilde{\mathrm{SL}(2)}$ est infini par des raisonnements généraux sans l'expliciter. En effet, étant donné que le groupe ondamental $\pi_1 G$ de tout groupe de Lie connexe G appartient au centre du revêtement simplement connexe $\widetilde{G}$, pour prouver que le centre de $\widetilde{\mathrm{SL}(2)}$ est infini, il suffit de montrer que le groupe $\pi_1\mathrm{SL}(2)$ est infini. On peut à cet effet se servir de la modification suivante de la proposition 8 de la leçon 12: si dans cette proposition le groupe G contient un sous-ensemble P (visiblement homéomorphe à la variété quotient G/H) tel que tout élément $g \in G$ se représente d'une seule façon sous la forme $g = hp$, où $h \in H$, $p \in P$, alors *le groupe* $\pi_1 H$ *est isomorphe au groupe* $\pi_1 G$. D'après le théorème de décomposition polaire, pour le sous-groupe $\mathrm{SO}(n)$ de $\mathrm{SL}(n)$, $n \geqslant 2$, ce sous-ensemble P sera l'ensemble des matrices unimodulaires définies positives d'ordre n. On démontre sans difficultés (pour $n = 2$ nous l'avons fait ci-dessus) que cet ensemble est homéomorphe à une boule ouverte d'un espace euclidien $\left(\text{de dimension } N = \frac{n(n+1)}{2} - 1\right)$ et, par suite, est simplement connexe. Donc, le groupe $\pi_1\mathrm{SL}(n)$ est isomorphe au groupe $\pi_1\,\mathrm{SO}\,(n)$ et, par conséquent, est un groupe cyclique infini pour $n = 2$. Nous avons opté pour une démonstration directe numérique dans la mesure où la forme explicite du groupe $\widetilde{\mathrm{SL}(2)}$ présente un intérêt intrinsèque.

Exposons maintenant l'origine géométrique des constructions algébriques formelles de la leçon 19.

On rappelle qu'une *forme différentielle* ω *de degré* m sur une variété

différentiable M est un champ (différentiable) de tenseurs antisymétriques de type $(m, 0)$. En coordonnées locales $x^1, \ldots, x^n$ sur M la forme ω s'exprime par la formule

$$\omega = f_{i_1 \ldots i_m} dx^{i_1} \wedge \ldots \wedge dx^{i_m},$$

où $f_{i_1 \ldots i_m}$ sont des fonctions différentiables sur le voisinage de coordonnées correspondant U, dépendant antisymétriquement des indices $i_1, \ldots, i_m$, et $dx^1, \ldots, dx^n$, des formes sur U de degré 1 dont les valeurs $(dx^1)_a, \ldots, (dx^n)_a$ en $a \in U$ (valeurs qui sont covecteurs de l'espace tangent $\mathsf{T}_a(M)$) forment une base duale de la base $\left(\frac{\partial}{\partial x^1}\right)_a, \ldots, \left(\frac{\partial}{\partial x^n}\right)_a$ de l'espace $\mathsf{T}_a(M)$.

L'ensemble $\Omega^m(M)$ des formes différentielles de degré $m \geqslant 0$ sur une variété M est un espace vectoriel (de dimension infinie) sur $\mathbb{R}$.

Les formes de degré 0 s'identifient canoniquement aux fonctions différentiables sur M et, par suite, l'espace vectoriel $\Omega^0(M)$, à l'algèbre $\mathcal{F}(M)$ des fonctions différentiables sur M.

La formule

$$(\varphi \wedge \psi)_a = \varphi_a \wedge \psi_a,$$

où a est un point quelconque de M, φ et ψ des formes différentielles sur M, définit une seule forme différentielle $\varphi \wedge \psi$ sur M, appelée *produit extérieur* des formes φ et ψ. Le degré de $\varphi \wedge \psi$ est égal à la somme des degrés de φ et de ψ.

La multiplication extérieure est associative, antisymétrique (si p et q sont les degrés des formes φ et ψ, alors $\psi \wedge \varphi = (-1)^{pq}\varphi \wedge \psi$) et est confondue avec la multiplication des fonctions pour les formes de degré 0.

Pour toute application différentiable $F: M \to N$, la formule

$$(F^*\omega)_a (A_1, \ldots, A_m) = \omega_{f(a)}((dF)_a A_1, \ldots, (dF)_a A_m),$$

où $a \in M$, $A_1, \ldots, A_m \in \mathsf{T}_a(M)$ et $\omega \in \Omega^m(N)$, définit sur M une forme différentielle $F^*\omega \in \Omega^m(M)$.

Pour toute fonction $f \in \mathcal{F}(M)$, la formule

$$(df)_a (A) = Af,$$

où $a \in M$ et $A \in \mathsf{T}_a(M)$, définit sur M une forme différentielle df de degré 1, appelée *différentielle* de la fonction f. En coordonnées locales

$$df = \frac{\partial f}{\partial x^i} dx^i.$$

L'opération d se prolonge de façon unique en une application linéaire (sur $\mathbb{R}$)

$$d: \Omega^m(M) \to \Omega^{m+1}(M),$$

définie pour tout $m \geqslant 0$ et possédant les trois propriétés suivantes:

1) l'application d est une différentiation pour le produit extérieur, c'est-à-dire que

$$d(\varphi \wedge \psi) = d\varphi \wedge \psi + (-1)^m \varphi \wedge d\psi,$$

quelles que soient les formes φ et ψ, où m est le degré de la forme φ;

2) $d \circ d = 0$,
c'est-à-dire que $d(d\omega) = 0$ pour toute forme ω;

3) pour toute application différentiable $F: M \to N$ et toute forme $\omega \in \Omega^m N)$, on a $dF^* \omega = F^* d\omega$.

De ces propriétés il résulte qu'en coordonnées locales l'opérateur d est défini par la formule

$$d\omega = df_{i_1 \ldots i_m} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_m}, \tag{2}$$

d'où son unicité. Pour prouver son existence, il faut prendre la formule (2) pour définition et démontrer qu'on obtient la même forme différentielle sur l'intersection de deux voisinages de coordonnées quelconques.

La forme $d\omega$ s'appelle *différentielle extérieure* de la forme ω.

On dit qu'une forme $\omega \in \Omega^m(M)$ est *fermée* si $d\omega = 0$, et *exacte*, s'il existe une forme $\varphi \in \Omega^{m-1}(M)$, telle que $d\varphi = \omega$. D'après la propriété 2) de l'opérateur d, l'espace vectoriel $B^m(M)$ des formes exactes est un sous-espace de l'espace $Z^m(M)$ des formes fermées, et l'on peut donc parler de l'espace quotient

$$H^m(M) = Z^m(M)/B^m(M).$$

Cet espace quotient s'appelle *espace* (ou *groupe*) *des cohomologies de de Rham* de la variété différentiable M.

Pour que le lien entre cette construction analytico-géométrique et les constructions algébriques formelles de la leçon 19 soit évident, on rappelle que le crochet de Lie munit l'espace vectoriel $\mathfrak{a}(M)$ des champs de vecteurs sur M d'une structure d'algèbre de Lie (de dimension infinie) et l'espace vectoriel $\mathcal{F}(M)$ des fonctions différentiables sur M, d'une structure de module sur l'algèbre de Lie $\mathfrak{a}(M)$ (par définition, $[X, Y]f = X(Yf) - Y(Xf)$ pour toute fonction $f \in \mathcal{F}(M)$). Nous sommes donc en droit de considérer l'espace vectoriel des cochaînes $C^m(\mathfrak{a}(M), \mathcal{F}(M))$ (dont la définition a un sens en dimension infinie).

Toute forme différentielle $\omega \in \Omega^m(M)$ définit au moyen de la

formule

$$u(X_1, \ldots, X_m)(a) = \omega_a((X_1)_a, \ldots, (X_m)_a), \quad a \in M,$$

où $X_1, \ldots, X_m$ sont des champs de vecteurs arbitraires sur M, une cochaîne $u \in C^m(\mathfrak{a}(M), \mathcal{F}(M))$. Il est évident que la correspondance $\omega \mapsto u$ est injective, donc les formes différentielles ω peuvent être identifiées aux cochaînes correspondantes (et désignées par les mêmes lettres).

Avec cette identification l'opérateur d sera défini par la formule

$$(3)\quad (m+1)(d\omega)(X_1, \ldots, X_{m+1}) = \\ = \sum_{i=1}^{m+1} (-1)^{i+1} X_i\omega(X_1, \ldots, \hat{X}_i, \ldots, X_{m+1}) + \\ + \sum_{i=1}^{m} \sum_{j=i+1}^{m+1} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{m+1}),$$

où $X_1, \ldots, X_{m+1}$ sont des champs arbitraires de $\mathfrak{a}(M)$, c'est-à-dire sera confondu (à un facteur multiplicatif inessentiel $m+1$ près) avec l'opérateur δ pour les cochaînes (le moyen le plus simple pour le prouver est de vérifier par un calcul direct que l'opérateur d défini par la formule (3) admet les propriétés 1) et 2) et est confondu avec la différentielle des fonctions auxquelles il est appliqué).

Ceci explique l'origine de l'opérateur δ. (A noter que toute cochaîne u ne se déduit pas automatiquement d'une forme différentielle: pour cela il est nécessaire et suffisant que pour tous champs de vecteurs $X_1, \ldots, X_m, Y_1, \ldots, Y_m$ et tout point $a \in M$, les égalités $(X_1)_a = (Y_1)_a, \ldots, (X_m)_a = (Y_m)_a$ entraînent l'égalité $u(X_1, \ldots, X_m)(a) = u(Y_1, \ldots, Y_m)(a)$. Donc, même si l'inclusion $\Omega^m(M) \subset C^m(\mathfrak{a}(M), \mathcal{F}(M))$ induit un certain homomorphisme

$$(4)\qquad H^m(M) \to H^m(\mathfrak{a}(M), \mathcal{F}(M)),$$

celui-ci ne sera pas en général un isomorphisme.)

On peut donner une interprétation topologique des espaces $H^m(\mathfrak{g}; \mathcal{V})$ pour le cas aussi où $\mathfrak{g}$ est une algèbre de Lie de dimension finie (sur $\mathbb{R}$). Pour simplifier, on se bornera au cas $\mathcal{V} = \mathbb{R}$. Une forme ω sur un groupe de Lie G est dite *forme invariante à gauche* si $L_a^*\omega = \omega$ pour tout élément $a \in G$.

Les formes invariantes à gauche de degré m forment un sous-espace $\Omega^m_{\text{inv}}(G)$ de l'espace vectoriel $\Omega^m(G)$, stable pour l'opérateur d (si $\omega = L_a^*\omega$, alors $d\omega = dL_a^*\,\omega = L_a^*d\omega$). Posons

$$H^m_{\text{inv}}(G) = Z^m_{\text{inv}}(G)/B^m_{\text{inv}}(G),$$

où $Z^m_{\mathrm{inv}}(G)$ est l'espace des formes invariantes à gauches fermées de degré m, et $B^m_{\mathrm{inv}}(G)$, son sous-espace qui est composé des formes $d\omega$, où $\omega \in \Omega^{m-1}_{\mathrm{inv}}(G)$.

Supposons comme toujours que e est l'unité du groupe de Lie G. L'espace vectoriel $\mathsf{T}_e G = \mathfrak{g}$ étant algèbre de Lie, les fonctionnelles multilinéaires antisymétriques sur $\mathsf{T}_e G$ ne sont autres que les cochaînes de cette algèbre de Lie sur le $\mathfrak{g}$-module trivial $\mathbb{R}$. Donc, la correspondance $\omega \mapsto \omega_e$ définit une application linéaire

$$\Omega^m_{\mathrm{inv}}(G) \to C^m(\mathfrak{g};\ \mathbb{R}). \tag{5}$$

Proposition 2. *L'application* (5) *est un isomorphisme.*

Démonstration. Une forme $\omega \in \Omega^m(G)$ est invariante à gauche (est élément de $\Omega^m_{\mathrm{inv}}(G)$) si et seulement si

$$\omega_a(A_1, \ldots, A_m) = u((dL_{a^{-1}})_a A_1, \ldots, (dL_{a^{-1}})_a A_m) \tag{6}$$

pour tout point $a \in G$ et tous vecteurs $A_1, \ldots, A_m \in \mathsf{T}_a(G)$, où $u = \omega_e$. Donc, si $\omega_e = 0$, alors $\omega = 0$.

Réciproquement, une vérification immédiate montre que pour toute cochaîne $u \in C^m(\mathfrak{g};\ \mathbb{R})$, la formule (6) définit une forme invariante à gauche $\omega \in \Omega^m_{\mathrm{inv}}(G)$ telle que $\omega_e = u$. □

Donc, les formes différentielles invariantes à gauche de $\Omega^m_{\mathrm{inv}}(G)$ peuvent être, en vertu de la proposition 2, identifiées aux cochaînes de $C^m(\mathfrak{g};\ \mathbb{R})$.

Par définition,

$$\omega_e(A_1, \ldots, A_m) = \omega(X_1, \ldots, X_m)(e)$$

pour toute forme $\omega \in \Omega^m(G)$ et tous vecteurs $A_1, \ldots, A_m \in \mathsf{T}_e G$, où $X_1, \ldots, X_m$ sont des champs de vecteurs arbitraires tels que $(X_1)_e = A_1, \ldots, (X_m)_e = A_m$. En particulier, sans nuire à la généralité, on peut admettre que les champs $X_1, \ldots, X_m$ sont invariants à gauche (sont éléments de l'algèbre de Lie $\mathfrak{g} = \mathfrak{l}(G)$).

Par ailleurs, si la forme ω est invariante à gauche, alors il est immédiat de voir que la fonction $\omega(X_1, \ldots, X_m)$ des champs de vecteurs invariants à gauche $X_1, \ldots, X_m$ est constante sur le groupe G, et, par suite,

$$Y\omega(X_1, \ldots, X_m) = 0$$

pour tout champ $Y \in \mathfrak{a}(G)$.

D'après la formule (3), on en déduit que pour toute forme $\omega \in \Omega^m_{\mathrm{inv}}(G)$, la valeur $(d\omega)_e(A_1, \ldots, A_{m+1})$ de la cochaîne $(d\omega)_e$ sur les vecteurs $A_1, \ldots, A_{m+1} \in \mathsf{T}_e G$ est donnée par la formule

$$(m+1)(d\omega)_e(A_1, \ldots, A_{m+1}) =$$
$$= \sum_{i=1}^{m} \sum_{j=i+1}^{m+1} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{m+1})(e),$$

où $X_1, \ldots, X_m$ sont des champs de vecteurs invariants à gauche sur G tels que $(X_1)_e = A_1, \ldots, (X_{m+1})_e = A_{m+1}$. En d'autres termes,

$$(m+1)(d\omega)_e(A_1, \ldots, A_{m+1}) =$$
$$= \sum_{i=1}^{m} \sum_{j=i+1}^{m+1} (-1)^{i+j} \omega_e([A_i, A_j], A_1, \ldots, \hat{A}_i, \ldots, \hat{A}_j, \ldots, A_{m+1}),$$

où $[A_i, A_j]$ est le crochet de Lie sur l'algèbre de Lie $\mathsf{T}_eG = \mathfrak{g}$ (c'est-à-dire que $[A_i, A_j] = [X_i, X_j]_e$).

Cela signifie que l'identification des formes invariantes à gauche de $\Omega^m_{\text{inv}}(G)$ aux cochaînes de $C^m(\mathfrak{g}, \mathbb{R})$ envoie l'opérateur d dans l'opérateur δ (à un facteur multiplicatif inessentiel $m+1$ près). Donc, pour tout $m \geqslant 0$

$$H^m_{\text{inv}}(G) = H^m(\mathfrak{g}; \mathbb{R}). \tag{7}$$

Contrairement au groupe $H^m(G)$, le groupe $H^m_{\text{inv}}(G)$ ne peut évidemment pas être traité comme un invariant topologique de la variété G_{diff}. Cependant, on peut démontrer que si un groupe de Lie est connexe et compact, les groupes $H^m(G)$ et $H^m_{\text{inv}}(G)$ sont isomorphes (ce qui en vertu de (7) nous donne entre autres une méthode assez efficace de calcul des groupes $H^m(G)$). Mais cette démonstration déborde le cadre de cet ouvrage.

LEÇON 21

Fonctionnelle de Killing d'un idéal.— Quelques propriétés des dérivations.— Radical et nilradical d'un idéal.— Prolongement des dérivations à une algèbre enveloppante universelle.— Idéaux d'une algèbre enveloppante de codimension finie.— Radical d'une algèbre associative.— Justification du pas inductif de la construction.— Démonstration du théorème d'Ado.— Conclusion.

Revenons maintenant à la démonstration du théorème d'Ado. Tous les éléments sont prêts, à l'exception de quelques lemmes purement techniques concernant essentiellement les dérivations et les algèbres enveloppantes.

Soit $\mathfrak{a}$ un idéal d'une algèbre de Lie $\mathfrak{g}$ et soit $\iota: \mathfrak{a} \to \mathfrak{g}$ une injection. Pour tout élément $x \in \mathfrak{a}$ l'opérateur linéaire $\operatorname{ad} \iota(x): \mathfrak{g} \to \mathfrak{g}$, tout en laissant $\mathfrak{a}$ invariant, induira des opérateurs $\mathfrak{g}/\mathfrak{a} \to \mathfrak{g}/\mathfrak{a}$ et $\mathfrak{a} \to \mathfrak{a}$, dont le premier est visiblement nul et le second, confondu avec $\operatorname{ad} x: \mathfrak{a} \to \mathfrak{a}$. Donc, pour tout couple $x, y \in \mathfrak{a}$, l'opérateur $\operatorname{ad} \iota(x) \circ \operatorname{ad} \iota(y)$ induira aussi dans $\mathfrak{g}/\mathfrak{a}$ un opérateur nul, et dans $\mathfrak{a}$ l'opérateur $\operatorname{ad} x \circ \operatorname{ad} y$. Donc, $\operatorname{Tr}(\operatorname{ad} \iota(x) \circ \operatorname{ad} \iota(y)) = \operatorname{Tr} (\operatorname{ad} x \circ$ $\circ \operatorname{ad} y)$, c'est-à-dire que $t_{\mathfrak{g}}(\iota(x), \iota(y)) = t_{\mathfrak{a}}(x, y)$.

Cela signifie que *la fonctionnelle de Killing $t_{\mathfrak{a}}$ de l'idéal $\mathfrak{a}$ est la restriction à $\mathfrak{a}$ de la fonctionnelle de Killing $t_{\mathfrak{g}}$ de l'algèbre $\mathfrak{g}$.*

Soit maintenant $D: \mathfrak{g} \to \mathfrak{g}$ une dérivation de l'algèbre de Lie $\mathfrak{g}$. Définissons sur la somme directe $\mathfrak{g}^* = \mathfrak{g} \oplus \mathsf{K}$ l'opération $[\,,]$ par la formule

$$[(x, \lambda), (y, \mu)] = ([x, y] + \lambda Dy - \mu Dx, 0),$$

$$x, y \in \mathfrak{g}, \qquad \lambda, \mu \in \mathsf{K}.$$

Il est évident que cette opération est bilinéaire et anticommutative. Par ailleurs, elle vérifie l'identité de Jacobi

$$[[(x, \lambda), (y, \mu)], (z, \nu)] + \\ + [[(y, \mu), (z, \nu)], (x, \lambda)] + [[(z, \nu), (x, \lambda)], (y, \mu)] = \\ = ([[x, y], z] + \lambda [Dy, z] - \mu [Dx, z] - \\ - \nu D([x, y] + \lambda Dy - \mu Dx), 0) + \\ + ([[y, z], x] + \mu [Dz, x] - \nu [Dy, x] - \\ - \lambda D([y, z] + \mu Dz - \nu Dy), 0) + \\ + ([[z, x], y] + \nu [Dx, y] - \lambda [Dz, y] - \\ - \mu D([z, x] + \nu Dx - \lambda Dz), 0) = 0.$$

Cette opération munit donc l'espace vectoriel $\mathfrak{g}^*$ d'une structure d'algèbre de Lie. En identifiant chaque élément $x \in \mathfrak{g}$ à l'élément $(x, 0) \in \mathfrak{g}^*$, on fait de $\mathfrak{g}$ un idéal de $\mathfrak{g}^*$. La dérivation intérieure ad ξ associée à l'élément $\xi = (0, 1)$ opère sur cet idéal au moyen de la formule

$$(\text{ad } \xi)x = [(0, 1), (x, 0)] = (Dx, 0) = Dx,$$

c'est-à-dire est confondue sur $\mathfrak{g}$ avec la dérivation D.

Par ailleurs, l'invariance de la fonctionnelle de Killing $t_{\mathfrak{g}^*}$ de l'algèbre de Lie $\mathfrak{g}^*$ appliquée aux éléments $x, y \in \mathfrak{g}$ et à ξ peut être exprimée sous la forme suivante:

$$t_{\mathfrak{g}^*}((\text{ad } \xi)x, \quad y) + t_{\mathfrak{g}^*}(x, (\text{ad } \xi)y) = 0,$$

d'où, d'après ce qui a été dit sur la fonctionnelle de Killing d'un idéal, il s'ensuit que *la fonctionnelle de Killing d'une algèbre de Lie* $\mathfrak{g}$ *vérifie la relation*

$$t_{\mathfrak{g}}(Dx, \ y) + t_{\mathfrak{g}}(x, Dy) = 0$$

qui exprime la propriété d'*invariance complète* de $t_{\mathfrak{g}}$

Un idéal $\mathfrak{a}$ d'une algèbre de Lie $\mathfrak{g}$ est dit *caractéristique* si $Dx \in \mathfrak{a}$ pour tout élément $x \in \mathfrak{a}$ et toute dérivation $D: \mathfrak{g} \to \mathfrak{g}$.

De la propriété d'invariance complète de $t_{\mathfrak{g}}$, on déduit que *l'annulateur* $\mathfrak{a}^{\perp}$ *relativement à* $t_{\mathfrak{g}}$ *d'un idéal caractéristique* $\mathfrak{a}$ *est idéal caractéristique*. En effet, si $x \in \mathfrak{a}^{\perp}$, alors $t_{\mathfrak{g}}(Dx, \ y) = -t_{\mathfrak{g}}(x, \ Dy) = 0$ pour toute dérivation D et tout élément $y \in \mathfrak{a}$, donc, $Dx \in \mathfrak{a}^{\perp}$. □

L'idéal $\mathfrak{g}^2$ étant visiblement caractéristique, on en déduit en vertu du corollaire 1 de la proposition 3 de la leçon 14 que *le radical* $\mathfrak{r}$ *de toute algèbre de Lie* $\mathfrak{g}$ *est idéal caractéristique*.

Puisque chaque dérivation intérieure d'une algèbre de Lie $\mathfrak{g}$ sur chacun de ses idéaux $\mathfrak{a}$ induit une dérivation (généralement pas

intérieure) de cet idéal, *tout idéal caractéristique* $\mathfrak{b}$ *de* $\mathfrak{a}$ *est idéal de l'algèbre* $\mathfrak{g}$ *tout entière.*

En particulier, le radical $\mathfrak{r}(\mathfrak{a})$ de l'idéal $\mathfrak{a}$ sera idéal de l'algèbre $\mathfrak{g}$. Etant idéal résoluble, ce radical est contenu dans le radical $\mathfrak{r} = \mathfrak{r}(\mathfrak{g})$ de l'algèbre $\mathfrak{g}$ et, par suite, dans l'intersection $\mathfrak{a} \cap \mathfrak{r}$. D'autre part, étant idéal résoluble de $\mathfrak{a}$, cette intersection est contenue dans $\mathfrak{r}(\mathfrak{a})$. Ceci prouve que *le radical* $\mathfrak{r}(\mathfrak{a})$ *d'un idéal* $\mathfrak{a}$ *d'une algèbre de Lie* $\mathfrak{g}$ *est confondu avec l'intersection de cet idéal avec le radical* $\mathfrak{r}$ *de l'algèbre tout entière*:

$$\mathfrak{r}(\mathfrak{a}) = \mathfrak{a} \cap \mathfrak{r}.$$

Considérons de nouveau l'algèbre de Lie $\mathfrak{g}^* \supset \mathfrak{g}$ construite à l'aide de la dérivation $D\colon \mathfrak{g} \to \mathfrak{g}$. Le radical $\mathfrak{r}$ de $\mathfrak{g}$ étant contenu dans le radical $\mathfrak{r}^*$ de $\mathfrak{g}^*$, il vient $D\mathfrak{r} = (\operatorname{ad} \xi)\mathfrak{r} \subset [\mathfrak{g}^*, \mathfrak{r}^*]$, et, par suite, $D\mathfrak{r} \subset [\mathfrak{g}^*, \mathfrak{r}^*] \cap \mathfrak{g}$.

Mais, d'après la proposition 4 de la leçon 16, l'idéal $[\mathfrak{g}^*, \mathfrak{r}^*]$ de $\mathfrak{g}^*$, donc l'idéal $[\mathfrak{g}^*, \mathfrak{r}^*] \cap \mathfrak{g}$ de $\mathfrak{g}$, sont nilpotents. Donc, l'idéal $[\mathfrak{g}^*, \mathfrak{r}^*] \cap \mathfrak{g}$ est contenu dans le nilradical $\mathfrak{n}$ de $\mathfrak{g}$, et, par suite, $D\mathfrak{r} \subset \mathfrak{n}$.

Formulons cette proposition sous forme de lemme afin de pouvoir s'y référer plus facilement:

Lemme 1. *L'image* $D\mathfrak{r}$ *du radical d'une algèbre de Lie* $\mathfrak{g}$ *par une dérivation* $D\colon \mathfrak{g} \to \mathfrak{g}$ *est contenue dans le nilradical* $\mathfrak{n}$ *de cette algèbre*:

$$D\mathfrak{r} \subset \mathfrak{n}. \ \square$$

Corollaire. *Le nilradical* $\mathfrak{n}(\mathfrak{a})$ *de tout idéal* $\mathfrak{a}$ *d'une algèbre de Lie* $\mathfrak{g}$ *est l'intersection de* $\mathfrak{a}$ *avec le nilradical* $\mathfrak{n} = \mathfrak{n}(\mathfrak{g})$ *de l'algèbre* $\mathfrak{g}$:

$$\mathfrak{n}(\mathfrak{a}) = \mathfrak{a} \cap \mathfrak{n}.$$

D é m o n s t r a t i o n. Soit $x \in \mathfrak{g}$. Alors, le lemme 1 appliqué à l'algèbre de Lie $\mathfrak{a}$ et à sa dérivation D induite par la dérivation intérieure $\operatorname{ad} x$ nous dit que $[x, \mathfrak{r}(\mathfrak{a})] \subset \mathfrak{n}(\mathfrak{a})$, donc que $[x, \mathfrak{n}(\mathfrak{a})] \subset \subset \mathfrak{n}(\mathfrak{a})$. Ceci signifie que la sous-algèbre nilpotente $\mathfrak{n}(\mathfrak{a})$ est idéal de $\mathfrak{g}$ et, par suite, est contenue dans $\mathfrak{n}$. Donc, $\mathfrak{n}(\mathfrak{a}) \subset \mathfrak{a} \cap \mathfrak{n}$. Comme l'inclusion inverse est évidente (l'intersection $\mathfrak{a} \cap \mathfrak{n}$ est idéal nilpotent de $\mathfrak{a}$), ceci prouve le corollaire. $\square$

Etudions maintenant les liens des dérivations et de l'algèbre enveloppante universelle $\mathfrak{U}$ de l'algèbre de Lie $\mathfrak{g}$ (cf. leçon 5).

Il est évident que chaque dérivation D d'une algèbre associative $\mathfrak{U}$ est également une dérivation de l'algèbre de Lie $[\mathfrak{U}]$ des commutateurs. Donc, si $D\mathfrak{g} \subset \mathfrak{g}$, D induit sur $\mathfrak{g}$ une dérivation $D\colon \mathfrak{g} \to \mathfrak{g}$. Peut-on obtenir ainsi toute dérivation $D\colon \mathfrak{g} \to \mathfrak{g}$ de l'algèbre de

Lie $\mathfrak{g}$, en d'autres termes, la dérivation $D: \mathfrak{g} \to \mathfrak{g}$ est-elle prolongeable en une dérivation $D: \mathfrak{U} \to \mathfrak{U}$ de l'algèbre associative $\mathfrak{U}$? La réponse est affirmative:

Lemme 2. *Toute dérivation D de l'algèbre de Lie $\mathfrak{g}$ se prolonge d'une seule manière en une dérivation (que l'on désignera par la même lettre D) d'une algèbre enveloppante universelle $\mathfrak{U}$.*

Démonstration. Etant donné que chaque élément de $\mathfrak{U}$ est un polynôme d'éléments de $\mathfrak{g}$, la dérivation $D: \mathfrak{U} \to \mathfrak{U}$, si elle existe, est définie de façon unique. Mais, comme les éléments de $\mathfrak{U}$ peuvent être généralement exprimés de plusieurs manières différentes par des éléments de $\mathfrak{g}$, la construction immédiate de la dérivation $D: \mathfrak{U} \to \mathfrak{U}$ implique une vérification assez laborieuse de sa cohérence. C'est pourquoi nous allons tourner cette difficulté en usant d'un artifice proposé par Jacobson.

Soit $\mathcal{A}$ l'algèbre des 2×2-matrices de la forme

$$\begin{pmatrix} u & v \\ 0 & u \end{pmatrix},$$

où $u, v \in \mathfrak{U}$. Considérons l'application

$$\varphi: x \mapsto \begin{pmatrix} x & Dx \\ 0 & x \end{pmatrix}$$

de l'algèbre de Lie $\mathfrak{g}$ dans l'algèbre $\mathcal{A}$, où D est une dérivation donnée de $\mathfrak{g}$. Comme

$$\begin{pmatrix} x & Dx \\ 0 & x \end{pmatrix}\begin{pmatrix} y & Dy \\ 0 & y \end{pmatrix} - \begin{pmatrix} y & Dy \\ 0 & y \end{pmatrix}\begin{pmatrix} x & Dx \\ 0 & x \end{pmatrix} =$$
$$= \begin{pmatrix} xy - yx & x\cdot Dy + Dx\cdot y - y\cdot Dx - Dy\cdot x \\ 0 & xy - yx \end{pmatrix} =$$
$$= \begin{pmatrix} [x, y] & [x, Dy] + [Dx, y] \\ 0 & [x, y] \end{pmatrix} = \begin{pmatrix} [x, y] & D[x, y] \\ 0 & [x, y] \end{pmatrix},$$

l'application φ est un homomorphisme de l'algèbre de Lie $\mathfrak{g}$ dans l'algèbre de Lie $[\mathcal{A}]$ des commutateurs de $\mathcal{A}$. L'algèbre $\mathfrak{U}$ étant universelle, il existe un homomorphisme $\psi = \mathfrak{U}\varphi$ de $\mathfrak{U}$ dans $\mathcal{A}$ tel que $\varphi = \psi \circ \iota$, où ι est l'injection $\mathfrak{g} \to \mathfrak{U}$. Du fait que $\mathfrak{U}$ est engendrée par les éléments de $\mathfrak{g}$, il s'ensuit immédiatement que pour tout élément $u \in \mathfrak{U}$, la matrice φu est de la forme

$$\varphi u = \begin{pmatrix} u & v \\ 0 & u \end{pmatrix}, \text{ où } v \in \mathfrak{U}.$$

Nous définissons une application $D: \mathfrak{U} \to \mathfrak{U}$ (visiblement linéaire) en posant $Du = v$.

Comme $\psi(uv) = \psi u \cdot \psi v$ pour tous éléments $u, v \in \mathfrak{U}$, c'est-à-dire que

$$\begin{pmatrix} uv & D(uv) \\ 0 & uv \end{pmatrix} = \begin{pmatrix} u & Du \\ 0 & u \end{pmatrix} \begin{pmatrix} v & Dv \\ 0 & v \end{pmatrix} = \begin{pmatrix} uv & u \cdot Dv + Du \cdot v \\ 0 & uv \end{pmatrix},$$

l'application $D: \mathfrak{U} \to \mathfrak{U}$ est une dérivation de l'algèbre $\mathfrak{U}$. Vu qu'elle prolonge visiblement la dérivation D de l'algèbre de Lie $\mathfrak{g}$, le lemme 2 est entièrement prouvé. □

Dans l'algèbre universelle $\mathfrak{U}$, nous porterons notre attention sur les idéaux (bilatères) $\mathcal{A}$ pour lesquels l'algèbre quotient $\mathfrak{U}/\mathcal{A}$ est de dimension finie. On dira que ces idéaux sont de codimension finie et on notera ceci codim $\mathcal{A} < \infty$.

On rappelle qu'un élément a d'une algèbre associative est *algébrique* s'il existe un polynôme $p(T)$ non nul tel que $p(a) = 0$.

On rappelle par ailleurs (cf. leçon 5), que pour toute base $x_1, \ldots, x_n$ d'une algèbre de Lie $\mathfrak{g}$ (que l'on supposera ici de dimension finie), les monômes de la forme $x_{i_1} x_{i_2} \ldots x_{i_s}$, où $i_1 \leqslant i_2 \leqslant \ldots \leqslant i_s$, i.e. de la forme $x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$, où $k_1 \geqslant 0, k_2 \geqslant 0, \ldots, k_n \geqslant 0$, forment une base de l'algèbre universelle $\mathfrak{U}$. De là on déduit sans peine que codim $\mathcal{A} < \infty$ *si et seulement si les classes* $\bar{x}_1 = x_1 + \mathcal{A}, \ldots, \bar{x}_n = x_n + \mathcal{A}$ *sont algébriques dans* $\mathfrak{U}/\mathcal{A}$. En effet, la condition d'algébricité de l'élément $\bar{x}_i$ signifie que l'élément $\bar{x}_i^{m_i}$ est une combinaison linéaire des éléments $1, \bar{x}_i, \ldots, \bar{x}_i^{m_i - 1}$ pour un certain $m_i \geqslant 1$. Donc, si les éléments $\bar{x}_1, \ldots, \bar{x}_n$ sont algébriques, tout monôme de la forme $\bar{x}_1^{k_1} \ldots \bar{x}_n^{k_n}$ (et, par suite, tout élément de l'algèbre quotient $\mathfrak{U}/\mathcal{A}$) est combinaison linéaire de monômes tels que $0 \leqslant k_i < m_i$ pour tout $i = 1, \ldots, n$. Les monômes vérifiant la dernière condition étant en nombre fini, ceci prouve que l'algèbre quotient $\mathfrak{U}/\mathcal{A}$ est de dimension finie, et, par suite, codim $\mathcal{A} < \infty$. La réciproque est évidente, puisque tout élément d'une algèbre de dimension finie est algébrique. □

A noter que la condition d'algébricité d'un élément $\bar{x}_i$ signifie qu'il existe un polynôme $p_i(T)$ non nul tel que $p_i(x_i) \in \mathcal{A}$.

On rappelle que le *produit* $\mathcal{A}\mathcal{B}$ de deux idéaux $\mathcal{A}$ et $\mathcal{B}$ d'une algèbre associative est, par définition, l'enveloppe linéaire des éléments ab, où $a \in \mathcal{A}$ et $b \in \mathcal{B}$. Ce produit est visiblement un idéal.

Lemme 3. *Le produit* $\mathcal{A}\mathcal{B}$ *de deux idéaux quelconques* $\mathcal{A}$ *et* $\mathcal{B}$ *de codimension finie est aussi un idéal de codimension finie.*

Démonstration. Si $p_i(T)$ et $q_i(T)$ sont des polynômes tels que $p_i(x_i) \in \mathcal{A}$ et $q_i(x_i) \in \mathcal{A}$, les polynômes $r_i(T) = p_i(T) \times q_i(T)$ sont tels que $r_i(x_i) \in \mathcal{A}$. □

Nous aurons besoin également de notions élémentaires sur les idéaux nilpotents dans les algèbres associatives.

Suivant la terminologie en usage en théorie des algèbres, on dit qu'un idéal $\mathcal{A}$ d'une algèbre associative $\mathcal{V}$ est ***nilpotent*** s'il existe un $k \geqslant 0$ tel que le produit de k quelconques de ses éléments est nul, c'est-à-dire si $\mathcal{A}^k$ est un idéal nul. Il est évident que la somme de deux idéaux nilpotents est un idéal nilpotent, d'où il résulte que dans toute algèbre associative $\mathcal{V}$ de dimension finie, il existe un plus grand idéal nilpotent $\mathcal{R}$ contenant tout autre idéal nilpotent. Cet idéal s'appelle *radical* de l'algèbre associative $\mathcal{V}$.

On dit qu'un idéal $\mathcal{A}$ d'une algèbre $\mathcal{V}$ est un *nilidéal* s'il est composé d'éléments nilpotents. Il est clair que tout idéal nilpotent est nilidéal. Il s'avère que, réciproquement, *tout nilidéal $\mathcal{A}$ d'une algèbre associative $\mathcal{V}$ de dimension finie est nilpotent* (et donc appartient à son radical $\mathcal{R}$). En effet, l'idéal $\mathcal{A}$ (ou, plus exactement, son algèbre $[\mathcal{A}]$ des commutateurs) étant sous-algèbre de l'algèbre de Lie $[\mathcal{V}]$ des commutateurs de $\mathcal{V}$, il est justiciable de la proposition 3 de la leçon 17. □

La somme $a + b$ de deux éléments nilpotents n'est généralement pas un élément nilpotent. La situation change si l'un des termes de cette somme appartient au radical.

Lemme 4. *Si $a \in \mathcal{R}$, et b est un élément nilpotent quelconque, la somme $a + b$ est nilpotente.*

Démonstration. Utilisons la proposition 3* de la leçon 17 (à noter que jusqu'ici on s'est contenté de la proposition 3) en prenant pour $\mathfrak{a}$ le sous-espace vectoriel de $\mathcal{V}$ engendré par le radical $\mathcal{R}$ et l'élément b, et pour $\mathfrak{g}$, la réunion de $\mathcal{R}$ et de b (puisque $[\mathcal{R}, b] \subset \mathcal{R}$, l'ensemble $\mathfrak{g}$ est stable pour la commutation). D'après cette proposition, le sous-espace $\mathfrak{a}$ est associativement nilpotent. Donc, en particulier, son élément $a + b$ est nilpotent. □

Dans le lemme suivant les références à la proposition 3* de la leçon 17 ne suffisent déjà plus et il faut revenir partiellement à sa démonstration.

Lemme 5. *Si une algèbre associative $\mathcal{V}$ de dimension finie est engendrée par une sous-algèbre $\mathfrak{h}$ de son algèbre des commutateurs $[\mathcal{V}]$, alors tout idéal $\mathfrak{n}$ de $\mathfrak{h}$ composé d'éléments nilpotents de $\mathcal{V}$ est contenu dans le radical $\mathcal{R}$ de $\mathcal{V}$.*

Démonstration. Il suffit de montrer que l'idéal $\mathcal{N}$ de $\mathcal{V}$ engendré par l'idéal $\mathfrak{n}$ est nilpotent. Comme $\mathfrak{h}$ engendre $\mathcal{V}$, tout élément de $\mathcal{N}$ est combinaison linéaire de produits d'éléments de $\mathfrak{n}$ et de $\mathfrak{h}$. Appelons *rang* d'un tel produit, le nombre de facteurs de $\mathfrak{n}$. Si un produit contient un facteur de la forme ax, $a \in \mathfrak{n}$ et $x \in \mathfrak{g}$, alors en se servant de la formule $ax = [a, x] + xa$, on peut le repré-

senter par une somme de deux produits de même rang dans lesquels le nombre de facteurs de $\mathfrak{g}$ soit est inférieur d'une unité, soit l'un de ces facteurs s'est déplacé à gauche. Donc, tout produit de rang r peut être représenté par une somme de produits de la forme ab, où a est un produit d'éléments de $\mathfrak{g}$ (éventuellement vide) et b, un produit de r éléments de $\mathfrak{n}$. Mais, d'après la proposition 3 de la leçon 16, la nilalgèbre de Lie $\mathfrak{n}$ est associativement nilpotente, c'est-à-dire qu'il existe un nombre n tel que tout produit de $r \geqslant n$ éléments de $\mathfrak{n}$ est nul. Donc, tout élément de l'idéal $\mathcal{N}$ qui est combinaison linéaire de produits de rang $\geqslant n$ est nul. Comme tout élément de l'idéal $\mathcal{N}^n \subset \mathcal{N}$ possède cette propriété par définition, ceci prouve que $\mathcal{N}^n = 0$, c'est-à-dire que l'idéal $\mathcal{N}$ est nilpotent. □

Nous pouvons entrer maintenant dans le vif de la démonstration du théorème d'Ado. Cette démonstration repose sur une construction inductive qui est possible grâce au lemme suivant:

Lemme 6. *Soient*

$\mathfrak{g}$ *une algèbre de Lie résoluble de dimension finie;*

$\mathfrak{n}$ *son nilradical;*

$\mathfrak{U}$ *une algèbre enveloppante universelle de* $\mathfrak{g}$;

$\mathcal{A}$ *un idéal de codimension finie de* $\mathfrak{U}$ *tel que pour tout élément* $x \in \mathfrak{n}$ *la classe* $x + \mathcal{A}$ *soit un élément nilpotent de l'algèbre quotient* $\mathfrak{U}/\mathcal{A}$.

Il existe alors un idéal $\mathcal{B}$ *de* $\mathfrak{U}$ *tel que:*

a) $\mathcal{B} \subset \mathcal{A}$;

b) codim $\mathcal{B} < \infty$;

c) *pour tout élément* $x \in \mathfrak{n}$, *la classe* $x + \mathcal{B}$ *est un élément nilpotent de l'algèbre quotient* $\mathfrak{U}/\mathcal{B}$;

d) *toute dérivation* D *de l'algèbre* $\mathfrak{U}$ *envoyant* $\mathfrak{g}$ *dans* $\mathfrak{g}$ *envoie* $\mathcal{B}$ *dans* $\mathcal{B}$.

D é m o n s t r a t i o n. Etant donné que tout homomorphisme d'algèbres associatives est également homomorphisme des algèbres des commutateurs respectives, l'image

$$\mathfrak{h} = (\mathfrak{g} + \mathcal{A})/\mathcal{A}$$

de l'algèbre de Lie $\mathfrak{g}$ dans l'algèbre $\mathcal{V} = \mathfrak{U}/\mathcal{A}$ est sous-algèbre de l'algèbre $[\mathcal{V}]$ des commutateurs engendrant $\mathcal{V}$. Comme par hypothèse l'image $(\mathfrak{n} + \mathcal{A})/\mathcal{A}$ de l'idéal $\mathfrak{n}$ dans $\mathcal{V}$ est composée d'éléments nilpotents, il vient, d'après le lemme 4, que cette image est contenue dans le radical $\mathcal{R}$ de $\mathcal{V}$. Donc, $\mathcal{R}$ contient aussi l'idéal de $\mathcal{V}$ engendré par cette image. Par conséquent, cet idéal est nilpotent. Le dernier idéal étant de la forme $\mathcal{C}/\mathcal{A}$, où $\mathcal{C}$ est l'idéal de $\mathfrak{U}$ engendré par $\mathfrak{n}$ et $\mathcal{A}$, ceci prouve qu'il existe un nombre r tel que l'idéal $\mathcal{B} = \mathcal{C}^r$ est contenu dans $\mathcal{A}$, c'est-à-dire possède la propriété a). D'après le lemme 3, l'idéal $\mathcal{B}$ possède la propriété b)

aussi. Comme pour tout $x \in \mathfrak{n}$, il existe par hypothèse un nombre $s \geqslant 0$ tel que $x^s \in \mathscr{A} \subset \mathscr{C}$, il s'ensuit que $x^{sr} \in \mathscr{C}^r = \mathscr{B}$ et, par suite, l'élément $x + \mathscr{B}$ de l'algèbre quotient $\mathfrak{U}/\mathscr{B}$ est nilpotent. Donc, l'idéal $\mathscr{B}$ possède aussi la propriété c).

Enfin, d'après le lemme 1, toute dérivation D de l'algèbre de Lie résoluble $\mathfrak{s}$ est telle que $D\mathfrak{s} \subset \mathfrak{n}$. Si donc la dérivation D de l'algèbre $\mathfrak{U}$ envoie $\mathfrak{s}$ dans $\mathfrak{s}$, c'est qu'elle envoie en fait $\mathfrak{s}$ dans $\mathfrak{n}$ et, par suite, l'algèbre $\mathfrak{U}$ tout entière dans son idéal engendré par $\mathfrak{n}$. Par conséquent, *a fortiori* $D\mathfrak{U} \subset \mathscr{C}$, et $D(\mathfrak{U}^r) \subset \mathscr{C}^r = \mathscr{B}$. Mais alors $D(\mathscr{B}) = D(\mathscr{C}^r) \subset D(\mathfrak{U}^r) \subset \mathscr{B}$, ce qui prouve la propriété d). □

On dira qu'une représentation ρ d'une algèbre de Lie $\mathfrak{g}$ est une *nilreprésentation* si pour tout élément $x \in \mathfrak{n}$ du nilradical $\mathfrak{n}$ de l'algèbre $\mathfrak{g}$, l'opérateur $\rho(x)$ est nilpotent.

Proposition 2 (pas inductif de la construction). *Si une algèbre de Lie $\mathfrak{g}$ se décompose en une somme directe*

$$\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{h}$$

d'un idéal résoluble $\mathfrak{s}$ et d'une sous-algèbre $\mathfrak{h}$, alors pour toute nilreprésentation σ de dimension finie de l'algèbre $\mathfrak{s}$, il existe une représentation ρ de l'algèbre $\mathfrak{g}$ telle que

$$\mathfrak{s} \cap \operatorname{Ker} \rho \subset \operatorname{Ker} \sigma.$$

Si l'algèbre de Lie $\mathfrak{g}$ est nilpotente ou, au contraire, si son nilradical est confondu avec celui de l'algèbre $\mathfrak{s}$, alors la représentation ρ peut être choisie parmi les nilreprésentations.

D é m o n s t r a t i o n. Considérons un élément arbitraire $x = y + z$, $y \in \mathfrak{s}$, $z \in \mathfrak{h}$, de l'algèbre $\mathfrak{g}$. Sa composante y définit au moyen de la formule L_y: $a \mapsto ya$, $a \in \mathfrak{U}$, une translation à gauche L_y: $\mathfrak{U} \to \mathfrak{U}$ de l'algèbre enveloppante universelle $\mathfrak{U}$ de l'algèbre de Lie $\mathfrak{s}$, et la composante z, la dérivation D_z de cette algèbre, qui est par définition le prolongement de la dérivation ad z de l'algèbre de Lie $\mathfrak{s}$. Définissons une application $\bar{\rho}$ de l'algèbre $\mathfrak{g}$ dans l'algèbre End $\mathfrak{U}$ des opérateurs linéaires $\mathfrak{U} \to \mathfrak{U}$ au moyen de la formule

$$\bar{\rho}(x) = L_y + D_z.$$

Soit $x_1 = y_1 + z_1$ un autre élément de l'algèbre de Lie $\mathfrak{g}$. Comme

$$[x, x_1] = [y, y_1] + [y, z_1] + [z, y_1] + [z, z_1] =$$
$$= ([y, y_1] - D_{z_1}y + D_z x_1) + [z, z_1],$$

il vient

$$\bar{\rho}([x, x_1]) = L_{[y, y_1]} - L_{D_{z_1}y} + L_{D_z x_1} + D_{[z, z_1]}.$$

Par ailleurs,

$$[\bar{\rho}(x), \bar{\rho}(x_1)] = [L_y + D_z, L_{y_1} + D_{z_1}] =$$
$$= [L_y, L_{y_1}] + [D_z, L_{y_1}] + [L_y, D_{z_1}] + [D_z, D_{z_1}].$$

Ceci étant,

$$[L_y, L_{y_1}] = L_y L_{y_1} - L_{y_1} L_y = L_{yy_1} - L_{y_1 y} = L_{[y, y_1]}.$$

Etant donné que sur $\mathfrak{s}$ la dérivation $D_{[z_1, z_2]}$ est confondue avec la dérivation $\mathrm{ad}[z, z_1] = [\mathrm{ad}\, z, \mathrm{ad}\, z_1]$ et les dérivations D_z et D_{z_1}, avec les dérivations $\mathrm{ad}\, z$ et $\mathrm{ad}\, z_1$, alors

$$[D_z, D_{z_1}] = D_{[z_1, z_2]}$$

sur $\mathfrak{s}$, donc sur $\mathfrak{U}$ tout entière.

Par ailleurs, étant donné que pour tout élément $a \in \mathfrak{U}$

$$[D_z, L_{y_1}]a = D_z L_{y_1} a - L_{y_1} D_z a =$$
$$= D_z(y_1 a) - y_1 D_z a = D_z y_1 \cdot a,$$

il vient

$$[D_z, L_{y_1}] = L_{D_z y_1}.$$

De façon analogue

$$[D_{z_1}, L_y] = L_{D_{z_1} y}.$$

Ceci prouve que

$$\bar{\rho}([x, x_1]) = [\bar{\rho}(x), \bar{\rho}(x_1)],$$

c'est-à-dire que $\bar{\rho}$ est un homomorphisme (« une représentation infinie ») de l'algèbre de Lie $\mathfrak{g}$ dans l'algèbre de Lie des opérateurs linéaires $\mathfrak{U} \to \mathfrak{U}$. A noter que cet homomorphisme est défini exclusivement par la décomposition $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{h}$ et ne dépend pas de la représentation σ.

Considérons maintenant une représentation σ. L'algèbre $\mathfrak{U}$ étant universelle, on peut prolonger cette représentation en un homomorphisme (que l'on désignera aussi par σ) de l'algèbre $\mathfrak{U}$ dans l'algèbre des opérateurs agissant dans l'espace de la représentation σ. La dernière algèbre étant de dimension finie, le noyau $\mathcal{A}$ de cet homomorphisme est de codimension finie. Puisque dire que la représentation σ est une nilreprésentation revient exactement à dire que pour tout élément x du nilradical $\mathfrak{n}$ de l'algèbre $\mathfrak{s}$, la classe $x + \mathcal{A}$ est un élément nilpotent de l'algèbre $\mathfrak{U}/\mathcal{A}$ (qui est isomorphe à l'algèbre linéaire $\sigma(\mathfrak{U})$), il vient que l'algèbre $\mathfrak{s}$ et l'idéal $\mathcal{A}$ sont justiciables du lemme 6. Donc, l'algèbre $\mathfrak{U}$ contient un idéal $\mathcal{B}$ possédant les propriétés a), b), c) et d) de ce lemme.

Comme $\mathcal{B}$ est un idéal, on a $L_y(\mathcal{B}) \subset \mathcal{B}$ pour tout élément $y \in \mathfrak{s}$ et puisque $\mathcal{B}$ possède la propriété d), il vient que $D_z(\mathcal{B}) \subset \mathcal{B}$ pour tout élément $z \in \mathfrak{h}$. Donc, $\bar{\rho}(x)(\mathcal{B}) \subset \mathcal{B}$ pour tout élément

$x \in \mathfrak{g}$ et, par suite, la formule

$$\rho(x)(u + \mathscr{B}) = \bar{\rho}(x)u + \mathscr{B}, \quad u \in \mathfrak{U},$$

définit un opérateur linéaire

$$\rho(x): \mathfrak{U}/\mathscr{B} \to \mathfrak{U}/\mathscr{B}.$$

L'application $\bar{\rho}: x \mapsto \bar{\rho}(x)$ étant un homomorphisme d'algèbres de Lie, l'application $\rho: x \to \rho(x)$ sera aussi un homomorphisme et donc, une représentation (car l'algèbre quotient $\mathfrak{U}/\mathscr{B}$ est de dimension finie en vertu de la propriété b)).

Si $x \in \mathfrak{s}$, alors $\rho(x) = L_{\bar{x}}$, où $\bar{x} = x + \mathscr{B}$. Donc, si $\rho(x) = 0$, alors $L_x(\mathfrak{U}) \subset \mathscr{B}$ et, par suite, $x \in \mathscr{B}$ (on rappelle que $\mathfrak{U}$ est une algèbre unitaire). Donc, $x \in \mathscr{A}$ (propriété a)) et, par suite, $\sigma(x) = 0$. Donc, $\mathfrak{s} \cap \operatorname{Ker} \rho \subset \operatorname{Ker} \sigma$.

Appliquons maintenant le lemme 5 à l'algèbre associative $\mathfrak{U}/\mathscr{B}$, à l'algèbre de Lie $(\mathfrak{s} + \mathscr{B})/\mathscr{B}$ et à son idéal $(\mathfrak{n} + \mathscr{B})/\mathscr{B}$. Ceci est possible, puisque, en vertu des propriétés b) et c), les conditions du lemme 5 sont remplies. Donc, en vertu de ce lemme, chaque élément $\bar{x} = x + \mathscr{B}$, $x \in \mathfrak{n}$, appartient au radical de l'algèbre associative $\mathfrak{U}/\mathscr{B}$ et, par suite, est nilpotent. Donc, l'opérateur linéaire $L_{\bar{x}}: \bar{a} \mapsto \bar{x}\bar{a}$ est nilpotent aussi. Comme $\rho(x) = L_{\bar{x}}$, ceci prouve que si le nilradical de l'algèbre $\mathfrak{g}$ est confondu avec le nilradical $\mathfrak{n}$ de l'algèbre $\mathfrak{s}$, la représentation ρ est une nilreprésentation.

Reste à traiter le cas où l'algèbre $\mathfrak{g}$ est nilpotente (et, par suite, en particulier, $\mathfrak{s} = \mathfrak{n}$). Nous devons prouver que dans ce cas l'opérateur $\rho(x)$ est nilpotent pour tout élément $x \in \mathfrak{g}$.

Supposons comme plus haut que $x = y + z$, où $y \in \mathfrak{s}$, $z \in \mathfrak{h}$. Comme $\mathfrak{s} = \mathfrak{n}$ dans ce cas, l'opérateur $\rho(y)$ est nilpotent d'après ce qui a été démontré ci-dessus. Bien plus, on peut formuler une proposition plus exacte en appliquant encore le lemme 5 à l'algèbre associative linéaire $\mathscr{T}$ engendrée par les opérateurs $\rho(x)$, $x \in \mathfrak{g}$. Cette algèbre est de dimension finie, est engendrée par l'algèbre de Lie $\rho(\mathfrak{g})$, quant à l'idéal $\rho(\mathfrak{n})$ de l'algèbre $\rho(\mathfrak{g})$, il est composé, comme nous venons tout juste de le voir, d'éléments nilpotents. Donc, d'après le lemme 5, les opérateurs $\rho(y)$, $y \in \mathfrak{n}$, appartiennent au radical $\mathscr{R}$ de l'algèbre $\mathscr{T}$.

Par ailleurs, l'algèbre $\mathfrak{g}$ étant nilpotente, tout opérateur ad z, $z \in \mathfrak{h}$, est nilpotent. Ceci signifie que toute dérivation $D_z: \mathfrak{U} \to \mathfrak{U}$ restreinte à $\mathfrak{s}$ est nilpotente. Puisque pour toute dérivation D, tout nombre N et tout couple d'éléments a et b, on a

$$D^N(ab) = \sum_{i=0}^{N} \binom{N}{i} D^i a \cdot D \quad ,$$

on déduit des relations $D^n a = 0$ et $D^m b = 0$ que $D^{n+m-1}(ab) = 0$. Donc pour tout produit $a \in \mathcal{U}$ d'éléments de $\mathfrak{z}$, il existe un $n(a)$ tel que $D_z^{n(a)}(a) = 0$. En passant à l'algèbre quotient $\mathcal{U}/\mathcal{B}$, on trouve que $\rho(z)^{n(a)}(\bar{a}) = 0$, où $\bar{a} = a + \mathcal{B}$. Comme l'algèbre $\mathcal{U}/\mathcal{B}$ est de dimension finie et possède une base $\bar{a}_1, \ldots, \bar{a}_s$ composée d'éléments de la forme $\bar{a}$, on en déduit qu'il existe un n tel que $\rho(z)^n(\bar{a}_i) = 0$ pour tout élément $\bar{a}_i$, $i = 1, \ldots, n$, de cette base. Mais alors $\rho(z)^n = 0$ sur $\mathcal{U}/\mathcal{B}$, c'est-à-dire que l'opérateur $\rho(z)$ est nilpotent.

Comme $\rho(x) = \rho(y) + \rho(z)$, pour achever la démonstration de la proposition 2, il reste à appliquer le lemme 4. □

Nous voilà enfin en mesure de prouver le théorème d'Ado sous une forme plus forte même:

Proposition 3 (théorème d'Ado). *Toute algèbre de Lie $\mathfrak{g}$ admet une nilreprésentation exacte.*

D é m o n s t r a t i o n. Construisons cette représentation par étapes.

P r e m i è r e é t a p e. Soit $\mathfrak{z}$ le centre d'une algèbre de Lie $\mathfrak{g}$, et soit $x_1, \ldots, x_m$ une base de $\mathfrak{z}$. Choisissons dans un espace vectoriel $\mathcal{V}$ (de dimension $\geqslant m + 1$) un opérateur nilpotent A tel que $A^{m+1} = 0$, mais $A^m \neq 0$, et posons

$$\rho_{\mathfrak{z}}(x) = \lambda_1 A + \ldots + \lambda_k A^k + \ldots + \lambda_m A^m$$

pour tout élément $x = \lambda_1 x_1 + \ldots + \lambda_k x_k + \ldots + \lambda_m x_m$ de $\mathfrak{z}$. Les opérateurs de la forme $\lambda_1 A + \ldots + \lambda_m A^m$ étant permutables et nilpotents, et de plus non nuls pour $(\lambda_1, \ldots, \lambda_m) \neq (0, \ldots, 0)$, l'application $\rho_{\mathfrak{z}}$ est une nilreprésentation exacte de l'algèbre $\mathfrak{z}$.

D e u x i è m e é t a p e. Soit $\mathfrak{n}$ le nilradical de l'algèbre de Lie $\mathfrak{g}$. Alors l'algèbre quotient $\mathfrak{n}' = \mathfrak{n}/\mathfrak{z}$ est nilpotente aussi et possède donc (cf. proposition 2 de la leçon 17) une suite d'idéaux

$$0 = \mathfrak{n}'_0 \subset \mathfrak{n}'_1 \subset \ldots \subset \mathfrak{n}'_m = \mathfrak{n}',$$

telle que $\dim \mathfrak{n}'_i = i$ pour tout $i = 0, 1, \ldots, m = \dim \mathfrak{n}'$ (dans la proposition 2 de la leçon 17 les inclusions sont inversées, mais cela est sans importance). Les images réciproques de ces idéaux dans $\mathfrak{n}$ forment une suite croissante d'idéaux

$$\mathfrak{z} = \mathfrak{n}_0 \subset \mathfrak{n}_1 \subset \ldots \subset \mathfrak{n}_m = \mathfrak{n},$$

commençant par le centre $\mathfrak{z}$ et telle que $\dim \mathfrak{n}_{i+1} = \dim \mathfrak{n}_i + 1$. Montrons par récurrence que pour tout $i = 0, 1, \ldots, m$, il existe une nilreprésentation ρ_i de l'idéal $\mathfrak{n}_i$ qui est exacte sur l'idéal $\mathfrak{n}_0 = \mathfrak{z}$.

Le premier pas de la récurrence n'est autre que la première étape

de la démonstration ($\rho_0 = \rho_{\mathfrak{z}}$). Supposons que pour un i on ait déjà construit une nilreprésentation ρ_i exacte sur $\mathfrak{z}$. En choisissant dans $\mathfrak{n}_{i+1}$ un sous-espace $\mathfrak{h}$ supplémentaire de $\mathfrak{n}_i$, et en remarquant que ce sous-espace est nécessairement algèbre de Lie, puisque de dimension un, on se place dans les conditions de la proposition 2 (avec $\mathfrak{s} = \mathfrak{n}_i$ et $\sigma = \rho_i$). L'idéal $\mathfrak{n}_{i+1}$ étant nilpotent, cette proposition nous dit qu'il existe une nilreprésentation ρ_{i+1} de l'idéal $\mathfrak{n}_{i+1}$ pour laquelle $\mathfrak{n}_i \cap \operatorname{Ker} \rho_{i+1} \subset \operatorname{Ker} \rho_i$ et qui de ce fait est encore exacte sur $\mathfrak{z}$.

Pour $i = m$, on obtient ainsi une nilreprésentation $\rho_{\mathfrak{n}} = \rho_m$ de l'idéal $\mathfrak{n}$ qui est exacte sur le centre $\mathfrak{z}$.

T r o i s i è m e   é t a p e. Soit $\mathfrak{r}$ le radical de l'algèbre de Lie $\mathfrak{g}$. En envisageant l'algèbre résoluble $\mathfrak{r}/\mathfrak{n}$ et en utilisant la proposition 1 de la leçon 16, on obtient, par la même construction que dans la deuxième étape, une suite de sous-algèbres

$$\mathfrak{n} = \mathfrak{r}_0 \subset \mathfrak{r}_1 \subset \ldots \subset \mathfrak{r}_m = \mathfrak{r},$$

commençant par $\mathfrak{n}$ et telle que pour tout $i = 0, 1, \ldots, m$, la sous-algèbre $\mathfrak{r}_i$ est un idéal de la sous-algèbre $\mathfrak{r}_{i+1}$. Nous pouvons appliquer encore la même procédure récurrentielle en commençant par la représentation $\rho_{\mathfrak{n}}$ du nilradical $\mathfrak{n}$ construit dans la deuxième étape. La proposition 2 peut maintenant être appliquée à chaque pas de la récurrence, car en vertu du corollaire du lemme 1, le nilradical de chaque algèbre $\mathfrak{r}_i$ est le même idéal $\mathfrak{n}$.

On obtient en définitive une nilreprésentation $\rho_{\mathfrak{r}}$ du radical $\mathfrak{r}$ exacte sur le centre $\mathfrak{z}$.

Q u a t r i è m e   é t a p e. Appliquons encore la proposition 2 à la décomposition de Lévi

$$\mathfrak{g} = \mathfrak{r} + \mathfrak{m}$$

de l'algèbre $\mathfrak{g}$ et à la nilreprésentation $\rho_{\mathfrak{r}}$. Comme $\mathfrak{n}(\mathfrak{r}) = \mathfrak{n}$, on obtient une nilreprésentation $\rho_{\mathfrak{g}}$ de l'algèbre de Lie $\mathfrak{g}$ exacte sur son centre $\mathfrak{z}$.

C i n q u i è m e   é t a p e. Considérons maintenant la représentation adjointe ad et sa somme directe

$$\rho = \operatorname{ad} \oplus \rho_{\mathfrak{g}}$$

avec la représentation $\rho_{\mathfrak{g}}$. Comme $\operatorname{Ker} \operatorname{ad} = \mathfrak{z}$ et $\mathfrak{z} \cap \operatorname{Ker} \rho_{\mathfrak{g}} = 0$, alors $\operatorname{Ker} \rho = 0$, c'est-à-dire que la représentation ρ est exacte. Comme la représentation ad est visiblement une nilreprésentation, et que la somme directe de deux nilreprésentations est manifestement nilreprésentation, ceci prouve entièrement la proposition 3. □

Ce n'est que maintenant qu'on peut tenir pour démontré le théorème de Cartan de la leçon 10 sur l'équivalence des catégories des groupes de Lie simplement connexes et des algèbres de Lie réelles! L'on comprend pourquoi nous avons trouvé la démonstration de ce théorème « peu satisfaisante » (l'autre démonstration du théorème de Cartan n'est pas plus simple, car elle repose aussi sur le théorème de Lévi). La démonstration directe du théorème de Cartan est donc une page qui reste à écrire dans l'histoire des mathématiques.

BIBLIOGRAPHIE

1. BOURBAKI N. — *Groupes et algèbres de Lie*. Chapitres 2 et 3. Hermann, Paris, 1972.
   Exposé détaillé. Difficile.

2. CHEVALLEV C. — *Theorie of Lie Groups*. Princeton University Press, 1946.
   Première et heureuse tentative d'exposition moderne de la théorie. En particulier, première élaboration soignée de la théorie des sous-fibrés intégrables (leçons 7 et 11) et construction d'une théorie des espaces de revêtement (leçon 8) n'utilisant pas les chemins.

3. DIXMIER J. — *Algèbres enveloppantes*. Gauthier-Villars, Paris, 1974.
   Le premier chapitre contient un exposé bien agencé mais très compendieux de nombreux points de théorie des algèbres de Lie. A cet ouvrage nous avons emprunté la démonstration du critère de Cartan de résolubilité (leçon 17).

4. DYNKIN E. — Algèbres de Lie normées et groupes analytiques (en russe). Uspekhi Mat. Nauk, 1950, 5, nº 1 (35), p. 135-186 (trad. anglaise: Amer. Math. Soc. Transl., (1) vol. 9, p. 470-534).
   Exposé détaillé des résultats de l'auteur concernant le calcul de la série de Campbell-Hausdorff (leçon 5). Contient aussi l'application de ces calculs à la construction d'un groupuscule de Lie par l'algèbre de Lie (leçon 6).

5. GLOUCHKOV V. — Structure des groupes localement bicompacts et cinquième problème de Hilbert (en russe). Uspekhi Mat. Nauk, 1957; 12, nº 2 (74), p. 3-41.
   Exposé détaillé de la résolution du cinquième problème de Hilbert.

6. JACOBSON N. — *Lie algebras*. Interscience Publishers, New-York, 1962.
   Monographie détaillée. Lecture difficile. Nous lui avons emprunté le schéma de la démonstration du théorème d'Ado (leçon 19).

7. KAPLANSKY I. — *Lie Algebras and Locally Compact Groups*. Chicago, 1971.
   Est composé de deux parties pratiquement indépendantes. La première contient un exposé des bases de la théorie des algèbres de Lie orienté vers la théorie générale des algèbres. Nous nous sommes inspiré de cet exposé dans la leçon 16. La deuxième partie est consacrée à la résolution du cinquième problème de Hilbert.

8. PONTRIAGUINE L. — *Groupes continus.* « Naouka », Moscou, 1973 (en russe).
Traité classique de la théorie des groupes différentiables et topologiques. S'agissant des groupes de Lie, il contient pratiquement tous les sujets de nos leçons; mais le troisième théorème de Lie (l'équivalence des catégories GR-LOC et ALG_f-LIE) a été démontré d'une autre manière (proche de celle de Lie), et le théorème de Cartan (l'équivalence des catégories GR_{00}-DIFF et ALG_f-LIE) a été établi en admettant la véracité du théorème de Lévi.

9. SERRE J. P. — *Algèbres de Lie. Semi-simples complexes.* New York, 1966.
Exposé assez détaillé et complet. Vise un lecteur averti.

10. TCHÉBOTAREV N. — *Théorie des groupes de Lie.* « Gostechizdat », Moscou, 1940 (en russe).
Première monographie en russe sur la théorie des groupes de Lie. Dans la lignée de Lie et de Cartan. Désuète, mais certains chapitres restent encore d'actualité.

BIBLIOGRAPHIE COMPLÉMENTAIRE

11. CANTOR I., SOLODOVNIKOV A. — *Nombres hypercomplexes.* « Naouka », Moscou, 1973 (en russe).
Introduction à la théorie des quaternions et des octonions contenant la démonstration des théorèmes de Hurwitz et de Frobenius.

12. FREUDENTHAL H. — *Oktaven, Ausnahmegruppen und Oktavengeometrie.* Mathematisch Instituut der Rigksunivrsiteit te Utrecht, 1951.
Exposé des recherches de l'auteur sur les groupes G_2, F_4 et E_6 et sur des problèmes connexes. Est destiné aux spécialistes. La lecture des articles est compliquée par une terminologie désuète et certaines particularités de style de l'auteur.

13. MASSEY W., STOLLINGS J. — *Algebraic topology. An introduction. Group theory and 3-manifolds.* London, 1971.
Exposé détaillé pour débutants de la théorie des espaces de revêtement et des groupes fondamentaux dans un « esprit classique », c'est-à-dire basé sur la notion d'homotopie des chemins continus.

INDEX DES MATIÈRES

